

BIOGEOCHEMICAL AND PHYLOGENETIC SIGNALS OF PROTEROZOIC AND PHANEROZOIC MICROBIAL METABOLISMS

by

Danielle S. Gruen

B.S., The Pennsylvania State University (2012)

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 2018

©2018 Danielle S. Gruen. All rights reserved.

The author hereby grants to MIT and WHOI permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document in whole or in
part in any medium now known or hereafter created.

Author

Joint Program in Applied Ocean Science & Engineering
Massachusetts Institute of Technology
& Woods Hole Oceanographic Institution
August 10, 2018

Certified by

Gregory Fournier
Assistant Professor
Massachusetts Institute of Technology
Thesis Supervisor

Accepted by

Shuhei Ono
Chair, Joint Committee for Chemical Oceanography
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution

BIOGEOCHEMICAL AND PHYLOGENETIC SIGNALS OF PROTEROZOIC AND PHANEROZOIC MICROBIAL METABOLISMS

by
Danielle S. Gruen

Submitted to the MIT-WHOI Joint Program in Oceanography and Applied Ocean Science and Engineering on August 10, 2018, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Oceanography

ABSTRACT

Life is ubiquitous in the environment and an important mediator of Earth's carbon cycle, but quantifying the contribution of microbial biomass and its metabolic fluxes is difficult, especially in spatially and temporally-remote environments. Microbes leave behind an often scarce, unidentifiable, or nonspecific record on geologic timescales. This thesis develops and employs novel geochemical and genetic approaches to illuminate diagnostic signals of microbial metabolisms. Field studies, laboratory cultures, and computational models explain how methanogens produce unique nonequilibrium methane clumped isotopologue ($^{13}\text{CH}_3\text{D}$) signals that do not correspond to growth temperature. Instead, $\Delta^{13}\text{CH}_3\text{D}$ values may be driven by enzymatic reactions common to all methanogens, the C-H bond inherited from substrate precursors including acetate and methanol, isotope exchange, or environmental processes such as methane oxidation. The phylogenetic relationship between substrate-specific methyl-corrinoid proteins provides insight into the evolutionary history of methylotrophic methanogenesis. The distribution of corrinoid proteins in methanogens and related bacteria suggests that these substrate-specific proteins evolved via a complex history of horizontal gene transfer (HGT), gene duplication, and loss. Furthermore, this work identifies a previously unrecognized HGT involving chitinases (ChiC/D) distributed between fungi and bacteria (~ 650 Ma). This HGT is used to tether fossil-calibrated ages from within fungi to bacterial lineages. Molecular clock analyses show that multiple clades of bacteria likely acquired chitinase homologs via HGT during the late Neoproterozoic into the early Paleozoic. These results also show that, following these HGT events, recipient terrestrial bacterial clades diversified ~ 400 - 500 Ma, consistent with established timescales of arthropod and plant terrestrialization. Divergence time estimates for bacterial lineages are broadly consistent with the dispersal of chitinase genes throughout the microbial world in direct response to the evolution and expansion of detrital-chitin producing groups including arthropods. These chitinases may aid in dating microbial lineages over geologic time and provide insight into an ecological shift from marine to terrestrial systems in the Proterozoic and Phanerozoic eons. Taken together, this thesis may be used to improve assessments of microbial activity in remote environments, and to enhance our understanding of the evolution of Earth's carbon cycle.

Thesis Supervisor: Gregory Fournier
Title: Assistant Professor
Massachusetts Institute of Technology

to my parents, Becky and Gary

ACKNOWLEDGMENTS

Although the personal stories are often excluded from peer review, science is nonetheless a human endeavor. The relationships formed during my PhD have taught me as many valuable lessons as the science itself. To everyone I have met and worked with over the past five years here, abroad, in my primary field of research and across many scientific disciplines, thank you. I am eternally grateful for the science, stories, jokes, technical training, tears, climbing beta, and inspiration that you have shared with me. To all of you, named and unnamed, thank you.

First, thank you to the professors who have generously advised me during my PhD, Greg Fournier and Shuhei Ono. Thank you for letting me into your world of science, sharing with me your expertise, and for ensuring that I leave MIT a stronger thinker and scientist. Shuhei, thank you for your support in all of our microbiology endeavors, from growing methanogens to sampling and analyzing microbial methane from all over the world. This led to an incredibly enriching project. Greg, thank you for your endless enthusiasm, willingness to take on any scientific challenge, and genuine support for your students. I am lucky to have had an advisor who gave me the freedom to explore my own interests as well as the resources and support to tackle some difficult scientific questions.

I feel fortunate to have been able to travel across Massachusetts and across the world to spend time with collaborators in other labs during my PhD. Thank you to the Holden Lab, Jim Holden, Lucy Stewart, Begum Topcuoglu, for kindly sharing your bench space, culturing techniques, and love for methanogens with me. You taught me that all forms of life grow better when you talk kindly to them, even (perhaps especially) microbes. Thank you to the Hristov Lab at Penn State for teaching me about your cows and allowing me to sample their methane. This served as the inspiration for our now beloved "hot" cow, Sally. Thank you to the Bohrmann Lab, especially to Gerhard Bohrmann and Thomas Pape, for sharing with me some of your most precious hydrates as well as your passion for these incredible environments. Finally, thank you to the Hinrichs Lab, especially Kai Hinrichs and Martin Koenneke, for supporting our Deep Life project, so warmly welcoming me to your lab, taking the time to work with me, and for allowing me to spend some of my favorite months of my PhD in Bremen learning, laughing, and culturing microbes with you. Your lab and university are filled with many impressive members from whom this work also benefited including Verena Heuer, Gunter Wegner, Heidi Taubner, Jenny Wendt, Xavier Prieto, Tobias Goldhammer, and Birgit Schmincke. Thank you also to the other Hinrichs Lab members and friends of the Lab including Nadine, Thomas, Flo, Xiaolei, Travis, Frauke, Rishi, Mark, Hendrik, Martin, Marcos, Felix, and Kevin. You all made the months I spent in Bremen, my days in the lab and evenings climbing in the Linie, something special.

Thank you also to the many MIT and WHOI professors and staff who provided valuable academic and personal guidance and facilitated my work and growth in the program, especially Meg Tivey and Bernhard Peucker-Ehrenbrink. Thank you to my committee, Greg Fournier, Shuhei Ono, and Stefan Sievert. Stefan, you served as both a wonderful microbiology professor and science mentor. Thank you to Jeff Seewald for kindly serving as my thesis proposal defense and thesis defense chair and providing optimism and guidance. My time in the joint program has benefited immensely from the administrative staff who go well above and beyond their job requirements to make things run smoothly for the students. Kris Kipp, Melody Abedinejad, and Roberta Allard, thank you (again and again) for everything

you have done for me over the years.

To my amazing JP peers, labmates, Oceans 13 group, and Building 54 and E25 friends, thank you to you all. I can't possibly name everyone, but know that each one of you has made an impact on my time here. Every day you impressed me, taught me, and supported me. In this regard, thank you especially to my cohort of JP Chem students: Cris, Gabi, Lauren, Net, Paul, and Tyler. You all have been one of the most supportive groups of people I have worked with, and this journey (and the months we spent studying in Clark) would not be nearly what it was without each one of you. A very special thank you to Cam for the balance, adventures, and support, especially at the end of this endeavor. Gaining academic siblings has been an unexpected benefit of doing a PhD. Andrew Whitehill and David Wang, thanks for teaching me, looking out for me, and growing up with me. Thank you also to Ono Lab members Bill, Eoghen, Shikma, Genming, and Jeemin. To the Fournier Lab: Jo, Abby, Cara, Kelsey, Lily, Makayla, Marjorie, Thiberio, Sarah, and Emilie it was a joy to work with you all. I'll miss our lab meetings and retreats. Abby, I am especially grateful for the pioneering code and pipelines you shared with us. Thank you also to my JP mentor, Kate French.

I am thankful to those who got me started on this path and to the people and communities who have deeply enriched my time as an MIT student, scientist, and human. Thank you to Chris House, The Penn State Astrobiology Research Center, and my Penn State Geosciences peers who became some of my first academic mentors and family. A big thanks also the MIT Ambulance and Cambridge/Somerville/BKB climbing communities for providing perspective, encouragement, and some invaluable life lessons during my time at MIT. A special thanks to Mark Forgues, Michael Beautyman, Lawrence Wong, the MIT HST GEMS program, and Scott Goldberg for important mentorship and for greatly enriching my PhD experience. Thanks also to Tyler, Julia, Brian, Gualtieri, Avi, Alex, Christine, Luis, and the climbing/lunch groups for pushing me to keep climbing and yet keeping me grounded through it all.

Finally, and most importantly, thank you to my parents and family who have provided unending support and encouragement. Rebecca and Gary, you have given me an insatiable curiosity for exploring, understanding, and improving this world. You have set the bar for what it means to achieve technical skill while remaining humbly dedicated to helping others. You have shown me how to deal with life's most dire emergencies and difficult obstacles with a smile, fierce optimism, and "focus." Every day I feel grateful to have you as parents, mentors, teachers, teammates, and friends. All that I achieve is indebted to you.

During my time in the MIT-WHOI Joint Program, I have been generously supported by the National Science Foundation (NSF), the NSF Graduate Research Fellowship Program, the MIT Energy Initiative and its partnership with Shell, the Neil and Anna Rasmussen Foundation Fund, and the Grayce B. Kerr Fellowship. This research and its dissemination was supported by funds from the Deep Carbon Observatory, NASA Astrobiology Institute, WHOI Academic Programs Office, and the MIT Graduate Student Council.

CONTENTS

1	Introduction	19
1.1	Motivation	20
1.2	Microbial Methanogenesis	21
1.3	Geobiological Approaches to Studying Biogeochemistry	23
1.3.1	Geochemical Approaches to Studying Microbial Biogeochemistry	23
1.3.2	Molecular Approaches to Studying Microbial Biogeochemistry	25
1.4	Thesis Overview	27
2	Experimental investigation on the controls of clumped isotopologue and hydrogen isotope ratios in microbial methane	29
2.1	Abstract	30
2.2	Introduction	30
2.3	Methods	33
2.3.1	Laboratory Culture Experiments	33
2.3.2	Sample preparation and Isotopologue Measurements	37
2.3.3	Isotope Notation and Calibration	37
2.3.4	FTIR Analysis of Methane Isotopologues	40
2.4	Results	42
2.4.1	Non-equilibrium $\Delta^{13}\text{CH}_3\text{D}$ signals of methane from microbial cultures	42
2.4.2	Effect of a closed system on $\delta^{13}\text{C}$, $\delta\text{D}_{\text{CH}_4}$ and $\Delta^{13}\text{CH}_3\text{D}$ systematics .	44
2.4.3	D-spiked H_2O experiments	44
2.4.4	D-labeled acetate experiment	46
2.5	Discussion	46
2.5.1	Closed system isotope effect does not explain non-equilibrium $\Delta^{13}\text{CH}_3\text{D}$	46
2.5.2	Origin of H in methane from three methanogenesis pathways	48
2.5.3	Pathway of D-isotope exchange during acetoclastic methanogenesis . .	51
2.5.4	Origin of ^{13}CH signals in methane from acetoclastic methanogenesis .	53
2.5.5	Non-equilibrium vs equilibrium $\Delta^{13}\text{CH}_3\text{D}$ signals of methane in the environment	57
2.6	Conclusions	58

3	Evolution of methylotrophic methanogenesis and possible implications for biogeochemical cycles in Proterozoic and Phanerozoic oceans	59
3.1	Abstract	60
3.2	Introduction	60
3.2.1	Microbial Methanogenesis	60
3.2.2	Horizontal Gene Transfer and Molecular Clock Dating	62
3.2.3	Microbial Links in Earth's Biogeochemical Cycles	64
3.2.4	Summary and Objectives	66
3.3	Methods	67
3.3.1	Identification of Genes of Interest and Test for Homology	67
3.3.2	Taxon Sampling	67
3.3.3	Sequences and Alignments	67
3.3.4	Phylogenetic Analyses and Model Selection	67
3.4	Results and Discussion	67
3.4.1	Single Gene Trees and Homologous Proteins	67
3.4.2	Homologous Corrinoid Proteins	68
3.4.3	Rooting the Methyl Corrinoid Tree	68
3.4.4	Corrinoid Protein Tree Topology and Substrate Specificity	68
3.4.5	Taxon Sampling: Opportunities and Limitations	73
3.4.6	Interrogating Microbial Carbon Metabolism in Marine Sediments with Methyl-Corrinoids	74
3.4.7	Outlook: Evaluating DMS Metabolism and the Timing of Methyl Specificity in Methanogenic and Bacterial Lineages	75
3.5	Conclusions	76
4	Chitinases: A standard candle for dating microbial lineages	119
4.1	Abstract	120
4.2	Background	120
4.2.1	Environmental Distribution of Chitin	121
4.2.2	Chitin Production and the Evolution of Fungi	122
4.2.3	The Evolution of Chitinase Gene Families	122
4.2.4	Bayesian Molecular Dating	123
4.2.5	Summary of Hypotheses and Objectives	126
4.3	Methods	126
4.3.1	Taxon Sampling	126
4.3.2	Sequences and Alignments	126
4.3.3	Phylogenetic Analyses	127
4.4	Results	128

4.4.1	Phylogeny of ChiD and ChiC homologs	128
4.4.2	Divergence Time Estimates of Bacterial Chitinases	128
4.4.3	Testing Molecular Clock Models	129
4.4.4	Impact of the Tree Process Prior and Rate Distribution Model	130
4.4.5	Impact of Taxon Sampling and Fungal Divergence Times	131
4.4.6	Impact of Calibrations	131
4.4.7	Informativeness of Sequence Data	132
4.5	Discussion	132
4.5.1	Fungal Origin and Distribution of Bacterial Chitinases	132
4.5.2	Importance of Chitinase Evolution for Dating Microbial Metabolisms .	133
4.5.3	Ecological Implications of Chitinase Evolution	134
4.6	Conclusion	135
4.7	Acknowledgements	135
5	Concluding Remarks and Future Directions	147
A	Nonequilibrium clumped isotope signals in microbial methane	151
B	Methane clumped isotopologue ($^{13}\text{CH}_3\text{D}$) measurements for determining the mixing of microbial and thermogenic methane in oceanic gas hydrates	161
C	The effect of the anaerobic oxidation of methane on clumped methane isotopologue ($^{13}\text{CH}_3\text{D}$) systematics	171
D	Evolution of the methyl-corrinoid pathway of methanogenesis and possible implications for biogeochemical cycling in Proterozoic and Phanerozoic oceans	179
E	Supplementary material for Chitinases: A standard candle for dating microbial lineages (Chapter 4)	193
F	Phylogenetic investigations of acetoclastic methanogenesis	205
G	References	211

LIST OF FIGURES

1-1	Genes and geochemistry are used to diagnose microbial metabolisms through time	22
1-2	Molecular evolution nomenclature	26
1-3	Topological conflicts between species and gene trees are used to infer HGT.	27
1-4	Pipeline for phylogenetic and molecular clock methods.	28
2-1	Clumped methane, $\Delta^{13}\text{CH}_3\text{D}$, plotted against environmental temperatures	41
2-2	Isotope systematics of $\delta^{13}\text{C}$, δD , $\Delta^{13}\text{CH}_3\text{D}$ over the course of a batch culture experiment	42
2-3	The δD and $\Delta^{13}\text{CH}_3\text{D}$ values of methane produced by <i>M. barkeri</i> in deuterium-spiked medium	43
2-4	Relative isotopologue abundances of methane produced by <i>M. barkeri</i> in batch cultures spiked with CD_3COO^-	44
2-5	The hydrogen isotope composition of microbial methane vs. the deuterium composition of media water in spiked culture experiments	45
2-6	The fractionation factor (α) as a function of the fraction of hydrogen derived from water (f) in acetoclastic methanogenesis in D_2O spiked water experiments	48
2-7	Three methanogenic pathways for <i>Methanosarcina</i>	49
2-8	Proposed mechanism of D-isotope exchange by acetate metabolism by <i>M. barkeri</i>	51
2-9	The values of $\Delta^{13}\text{CH}_3\text{D}$ (or Δ_{18}) of methane plotted against D/H fractionation factor between methane and water for methanogen cultures grown on different substrate	55
3-1	Metanogenesis-relevant timeline of events in the evolution of the carbon cycle	78
3-2	Methanogenesis Pathways	79
3-3	Pathway of Methylophilic Methanogenesis	80
3-4	Simplified Methyl-Corrinoid Pathway	81
3-5	A presence-absence diagram depicting both the phylogenetic relationships between taxa on the species tree and the genes associated with each pathway	82
3-6	Relationship between major mediators of DMSP production and degradation in the marine environment	83
3-7	Gene tree of methyl-corrinoid (mtaC) homologs	84

4-1	Bayesian Molecular Dating Parameters	124
4-2	Chitinase Gene Tree	136
4-3	Chronogram depicting the phylogenetic relationships between chitinase taxa and posterior age estimates obtained under the prior	137
4-4	Posterior chitinase date distribution across nodes with varied model selection	138
4-5	Posterior chitinase date distribution across nodes with varied calibration . . .	139
4-6	Prior date distributions across nodes	140
4-7	Prior vs. posterior chitinase date distributions across nodes.	141
A-1	Isotopologue compositions of methane samples	156
A-2	Extent of clumped- and hydrogen-isotopic disequilibria in methane	157
A-3	$\Delta^{13}\text{CH}_3\text{D}$ values of methane produced by hydrogenotrophic methanogens in batch cultures reflect kinetic effects	158
A-4	Relationships between $\Delta^{13}\text{CH}_3\text{D}$ and H_2 concentration for microbial methane	159
B-1	Schematic of methane hydrate formation	163
B-2	Hypothesized relationship between sediment depth and clumped methane. . .	164
B-3	Sample site locations.	165
B-4	Three sample sites highlighted	166
B-5	Depth profile of $\delta^{13}\text{C}$, δD , and $\Delta^{13}\text{CH}_3\text{D}$ at the Pechori Mound site, Eastern Black Sea.	167
B-6	Mixing Model illustrating the relationship between the mixing of two sources of $\delta^{13}\text{C}$, δD , and highlighting the nonlinearity of mixing $\Delta^{13}\text{CH}_3\text{D}$ of two sources.	167
C-1	$\delta^{13}\text{C}$ as a function of methane consumed via AOM.	173
C-2	δD as a function of methane consumed via AOM.	174
C-3	$\Delta^{13}\text{CH}_3\text{D}$ as a function of methane consumed via AOM.	175
C-4	$\Delta^{13}\text{CH}_3\text{D}$ -based temperature as a function of methane consumed via AOM. .	176
D-1	Gene tree of cmtA	181
D-2	Gene tree of mtaA	182
D-3	Gene tree of mtaB	183
D-4	Gene tree of mtbA	184
D-5	Gene tree of mtbB	185
D-6	Gene tree of mtbB1	186
D-7	Gene tree of mtbC	187
D-8	Gene tree of mtmB1	188
D-9	Gene tree of mtsA	189
D-10	Gene tree of mtsB	190

D-11	Gene tree of mttb	191
D-12	Gene tree of mttC	192
E-1	Expanded RAxML gene tree with tip labels and bootstrap support values. . .	194
E-2	Prior date distributions across nodes under the four calibration setups. . . .	195
E-3	Chronogram with corresponding node numbers used in analysis.	196
F-1	The relationship between donor and recipient lineages of acetoclastic methanogens	208
F-2	Concatenated AckA Pta gene tree of methanogens and their closest relatives.	209

LIST OF TABLES

2-1	Summary of experiments.	34
2-2	Results for methanogen culture experiments.	36
2-3	Isotopic compositions of substrates and medium water.	40
3-1	Methyltransferase pathway genes of interest	85
3-2	DMS Genes of Interest.	86
3-3	DMSP sources, relevant organisms, and proposed evolutionary dates.	87
3-4	Taxa list	88
4-1	Phylobayes model parameters tested in this study	142
4-2	Calibrations used in molecular clock models	143
4-3	Posterior divergence time estimates calculated under model assumptions . . .	144
4-4	Meta-analysis of posterior date distributions for model parameters	145
B-1	Sample descriptions and associated references.	168
B-2	Methane hydrate data. $\delta^{13}\text{C}$ is in reference to PDB, δD is in reference to SMOW. $\delta^{13}\text{C}$ (B) indicates that this was measured in Bremen.	169
C-1	Methane isotope composition over the course of the methane consumption experiment.	177
E-1	Taxa in this study	197
E-2	Extended table of all results and parameters tested	203

CHAPTER 1

INTRODUCTION

1.1 Motivation

Today, the Earth’s surface, ocean, and deep sediments are occupied by life and its metabolic products (Nealson and Conrad, 1999). Plants, bacteria, fungi, archaea, and arthropods dominate the Earth’s biomass (Bar-On et al., 2018). The evolution of these carbon-bearing organisms has thus dramatically affected the distribution of organic molecules and fluxes in Earth’s biogeochemical cycles (Schirrmeister et al., 2013; Hinrichs, 2002; Shih, 2015; Lyu et al., 2018; Kenrick et al., 2012; Falkowski et al., 2008). Quantifying the contribution of microbes in the modern environment, both in terms of their organic biomass as well as their metabolic activities, remains a significant challenge (Bar-On et al., 2018). This challenge becomes even greater when we consider how life has evolved and shaped its environment on geologic timescales (Knoll, 2017; Knoll and Nowak, 2017). Understanding this interplay between life and Earth requires a way to estimate the impact these processes have today and to date when these processes arose. Developing proxies and analytical tools to understand the signals associated with some of the simplest biopolymers from the biomass, enzymatic activity, and metabolic byproducts of key taxonomic groups may enable us to better date when these processes became environmentally relevant and how they have impacted Earth’s biogeochemistry.

In the absence of direct biological or physical evidence, we rely on proxies of microbial processes in geographically or temporally extreme environments as diagnostic markers of the unique role microbes play in the Earth system (Peckmann and Thiel, 2004; Hinrichs, 2002; Braakman et al., 2017; Bontognali et al., 2012; Donoghue and Benton, 2007). Figure 1-1 illustrates some of what we know about the evolution of Earth’s carbon cycle from such proxies. New technologies have afforded us the ability to make increasingly precise measurements of stable isotopes of carbon and hydrogen, which provide clues of formation processes; however, the specific microbial mechanisms that generate observed isotopic signatures are often difficult to interpret (Wang et al., 2015; Eiler et al., 2014; Stolper et al., 2013, 2014; Douglas et al., 2017; Young et al., 2017). Despite progress in understanding the molecular record of extant bacterial genomes in the modern environment, the timing of the evolution of major clades of microbial life remains challenging to resolve (Dos Reis et al., 2015) due to complex gene histories and a lack of clear morphological traits (Shih, 2015). In the absence of physical (geochemical or fossil) records of microbial evolution, we rely on genetic clues to interpret and time the evolutionary history of microbial lineages (Schirrmeister et al., 2015; Parfrey et al., 2011). This thesis aims to unscramble the sometimes obscure geochemical and phylogenetic environmental imprints of microbial metabolisms and to understand the mechanisms that may drive these signals.

1.2 Microbial Methanogenesis

Methane is a key species in the global carbon cycle (Alperin and Blair, 1992), a potent greenhouse gas (Wecht et al., 2014), a source of energy (Whiticar, 1990), and a potential biosignature (Webster et al., 2014). Over the last four billion years, Earth has shifted from a dominantly abiogenic methane world to a dominantly microbial methane world (Figure 1-1) (Whiticar, 1990; Whalen, 2005). This shift has coincided with major events in the evolution of both life and Earth (Rothman et al., 2014). Historically, building a methane budget to understand the major sources and sinks of methane has proven challenging (Wang et al., 2015). This is driven, in part, by a temporal and spatial disconnect between our sampling abilities and environmentally-significant reservoirs of methane, as many are too remote in space or time to sample directly (Orcutt et al., 2013; Wolfe and Fournier, 2018). In addition, the organisms that interact with carbon cycle intermediates may be hard to identify (Ijiri et al., 2018) or be part of cryptic metabolic processes (Holmkvist et al., 2011).

The greatest source of methane today is produced by methanogenic microbes in anoxic environments such as swamps, sediments, rice paddies, and ruminant tracts via methanogenesis (Klapp et al., 2010). Microbial methanogenesis also forms the majority of gas trapped in the largest reservoir of methane on Earth: oceanic gas hydrates (Thauer et al., 2008; Kvenvolden, 1993). Anaerobic archaea are the only organisms capable of producing methane as a catabolic end product, via three pathways specific to certain classes of substrates. Hydrogenotrophic methanogenesis (1) is the reduction of CO_2 with H_2 as an electron donor (Baptiste et al., 2005). Acetoclastic methanogenesis (2) use of acetate as a terminal electron acceptor. Methylo-trophic methanogenesis (3) uses methyl compounds (methanol, methyl amines, dimethylsulfide, or methylthiols) as substrates (Penger et al., 2012; Baptiste et al., 2005). The distribution of methanogens that carry out these three metabolisms is an active area of research, and putative methanogens with the ability to use a diverse set of substrates are still being discovered (Orcutt et al., 2011, 2013; Vanwonterghem et al., 2016; Thauer et al., 2008).

As one of the earliest metabolic pathways to evolve (Baptiste et al., 2005), microbial methanogenesis has played an important role in the biogeochemical cycling of carbon not only in the modern environment, but also since the evolution of the methanogenic metabolism (Rothman et al., 2014) over 3.5 billion years ago (Ga) (Ueno et al., 2006). Methanogenesis has remained one of the most highly conserved metabolic strategies, and, interestingly, the pathway has never been transferred to another group of microbes (Fournier et al., 2009). Nonetheless, this metabolism has continued to evolve in ways that increase its metabolic capabilities and impact major biogeochemical cycles (Evans et al., 2015). For example, it is hypothesized that acetoclastic methanogens acquired the ability to use the substrate acetate from a group of cellulytic bacteria, *Clostridia* (Fournier et al., 2009; Rothman

et al., 2014). The sometimes geographically remote environments that harbor methanogenesis and other globally-relevant metabolisms are also often some of the most interesting and informative pieces to Earth's biogeochemical puzzle. Thus, novel geochemical and phylogenetic approaches are needed to illuminate these hard to reach places.

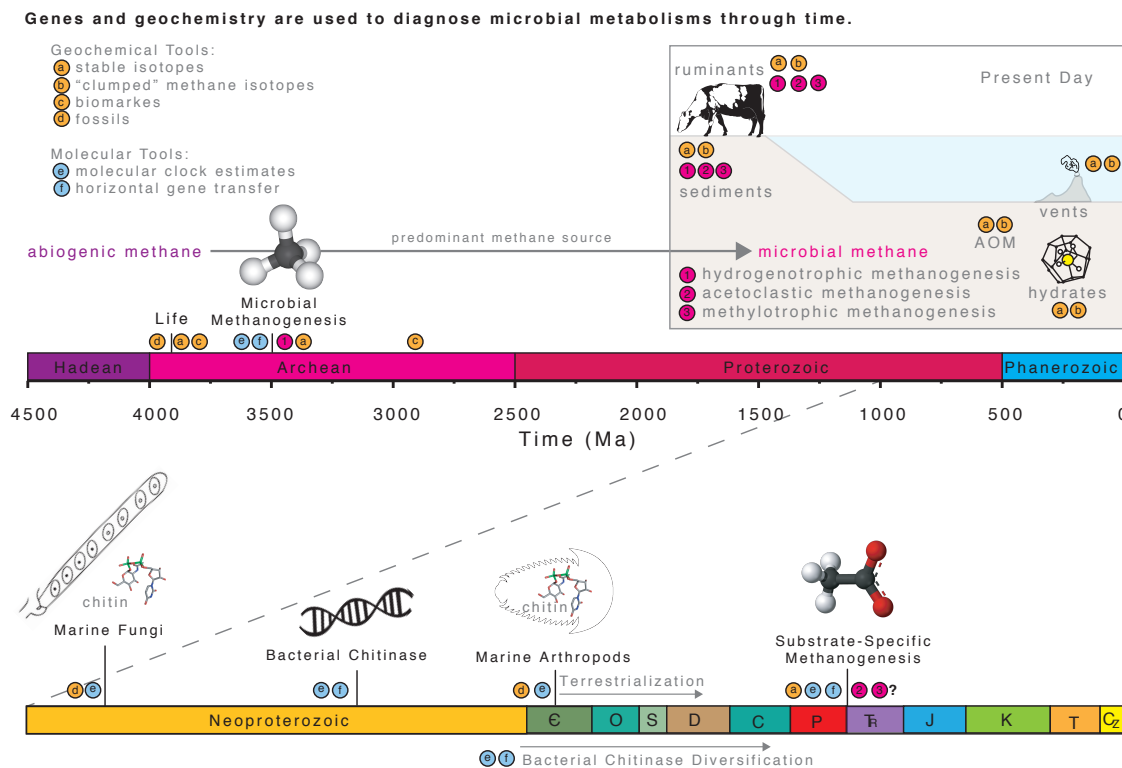


Figure 1-1: Genes and geochemistry are used to diagnose microbial metabolisms through time. This timeline highlights some of the key biologic innovations that have affected the carbon cycle discussed in this thesis. The timeline is populated by events that are informed by both geochemical and phylogenetic tools. The middle bar illustrates 4.5 billion years of Earth history. This bar is delineated by zircons (Wilde et al., 2001), the origin of life (Knoll, 2017), and microbial methanogenesis corresponding to 3.5 Ga molecular clock estimates (Wolfe and Fournier, 2018), 3.46 Ga fluid inclusions (Ueno et al., 2006), and 2.7 Ga isotopically depleted kerogen (Hinrichs, 2002). Some present-day sources sampled and described in this thesis are illustrated in the upper right box. The lower bar zooms into part of the Proterozoic and Phanerozoic Eons, and highlights the evolution of key marine and terrestrial eukaryotic, bacterial, and archaeal organisms that have interacted with the carbon cycle. This includes the emergence of organisms that produce one of Earth's most abundant biopolymers: chitin (Lozano-Fernandez et al., 2016; Berbee et al., 2017; Wolfe et al., 2016; Daley et al., 2018; Floudas et al., 2012; Sharpe et al., 2015). Substrate-specific acetoclastic methanogenesis is defined by ~2.4 Ma phylogenetic estimates (Rothman et al., 2014). Substrate-specific pathways of methylotrophic methanogenesis are hypothesized to have evolved in the Proterozoic or Phanerozoic. Because this pathway has never been dated, it appears with a question mark.

1.3 Geobiological Approaches to Studying Biogeochemistry

Defining and quantifying the role of microbial metabolism in Earth’s carbon biogeochemistry is important not only for attributing the distribution of elements on Earth over the last 4.5 billion years and how Earth-life processes have evolved, but also for predicting how these systems may be affected by future changes (Whalen, 2005; McCalley et al., 2014). Moreover, understanding the carbon cycle both in the past and today relies on an understanding of Earth’s geochemical signals as well as life’s biological signals (Knoll, 2017; Schirrmeister et al., 2015). Taking a geologic perspective to studying biology introduces significant challenges, but it also enables the application of a wide set of interdisciplinary tools to understand those environments that cannot be sampled directly (Knoll, 2014).

1.3.1 Geochemical Approaches to Studying Microbial Biogeochemistry

Fossils and Biomarkers

Diagnostic fossils are one of the most informative pieces of evidence we have to identify what lived in the past. For instance, the fossil record of fungi provides evidence of some of the oldest animal life (Berbee et al., 2017). More specifically, crown Ascomycetes are used to inform the minimum age of one of the major splits in fungal evolution (Berbee and Taylor, 2010). Fossil evidence is also used to provide a minimum age estimate for the oldest crown Arthropods, which likely originated in the Cambrian (Lozano-Fernandez et al., 2016; Wolfe et al., 2016; Daley et al., 2018). When fossils do not exist for taxa of interest (*e.g.*, we are looking too far back in time or the organisms lack clear morphologies Shih, 2015), biomarkers, a form of chemical fossil, can provide additional metabolic information (*e.g.*, Peters et al., 2005; Zhuang et al., 2016; Inagaki et al., 2015; Brocks and Pearson, 2005).

Carbon and Hydrogen Stable Isotopes of Methane

Carbon ($^{13}\text{C}/^{12}\text{C}$) and hydrogen (D/H) isotope ratios are widely used to identify the source of environmental methane (thermogenic, microbial, or abiogenic) both today and over geologic time (*e.g.*, Blair and Carter, 1992; Whiticar, 1999; Conrad et al., 2009; Blaser and Conrad, 2016; Hinrichs, 2002; Ueno et al., 2006). This is based on the observation that the enzymatic reactions carried out by microbes, for example, preferentially incorporate the "lighter" isotopes, resulting in isotopically depleted isotope values (Hayes, 2001). The notation for describing carbon and hydrogen stable isotopes is delta (δ), and is reported in units of permil (‰):

$$\delta^{13}\text{C} = \frac{(^{13}\text{C}/^{12}\text{C})_{\text{sample}}}{(^{13}\text{C}/^{12}\text{C})_{\text{PDB}}} - 1 \quad (1.1)$$

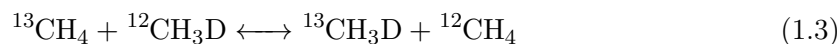
$$\delta D = \frac{(D/H)_{\text{sample}}}{(D/H)_{\text{SMOW}}} - 1 \quad (1.2)$$

where, PDB and SMOW are Pee Dee Belemnite and Standard Mean Ocean Water, respectively.

The isotopic composition of methane is mediated by both kinetic and equilibrium isotope effects during formation processes (Valentine et al., 2004; Whiticar, 1990). Microbial enzymatic reactions are typically associated with a general range of fractionation factors and more depleted environmental isotopic compositions (Summons et al., 1998; Whiticar, 1999; Valentine et al., 2004; Londry et al., 2008; Conrad et al., 2009). However, microbes can also produce highly variable and often overlapping carbon and hydrogen isotope signals, affected by the source of carbon and hydrogen as well as isotope fractionation during formation (Valentine et al., 2004; Penger et al., 2012; Alperin and Blair, 1992; Sugimoto and Wada, 1993; Krzycki and Kenealy, 1987; Waldron et al., 1999; Yoshioka et al., 2008; Kawagucci et al., 2014). Thus, carbon and hydrogen isotope values alone cannot always unambiguously determine the methane source.

Methane "Clumped" Isotopologues

It was proposed that clumped isotopologues of methane may avoid some of the challenges inherent to identifying methane sources based on conventional carbon and hydrogen stable isotopes, because clumped isotopologue formation is mediated by equilibrium processes dependent on methane formation temperature (Stolper et al., 2013, 2014, 2015; Eiler et al., 2014; Wang et al., 2015). 'Clumped isotopologue' refers to the multiply substituted isotopologues of methane; those molecules that have more than one rare isotopic constituent clumped together in a single molecule (*e.g.*, $^{13}\text{CH}_3\text{D}$). The following reaction shows equilibrium among four methane isotopologues, including $^{13}\text{CH}_3\text{D}$:



The equilibrium constant (K_T) can be written as:

$$K_T = \frac{[^{13}\text{CH}_3\text{D}][^{12}\text{CH}_4]}{[^{12}\text{CH}_3\text{D}][^{13}\text{CH}_4]} \quad (1.4)$$

The value of K_T primarily depends on temperature, and it approaches unity at high temperatures (1.0002 at 1,000°C), but is about 1.0057 at 25°C as estimated by molecular simulations (Ma et al., 2008; Stolper et al., 2015; Wang et al., 2015; Liu and Liu, 2016). Thus, the precise measurements of four isotopologues' abundance ($^{12}\text{CH}_4$, $^{13}\text{CH}_4$, $^{12}\text{CH}_3\text{D}$, and $^{13}\text{CH}_3\text{D}$) were thought to provide an estimate of the temperature at which the methane gas was formed or thermally equilibrated. Initial studies using high-resolution

mass-spectrometry demonstrated that this new isotopologue thermometer provides a range of temperatures that are consistent with formation temperatures for methane samples from geologic environments, such as natural gas reservoirs (Stolper et al., 2014). Later studies, however, showed that methane sampled from surface environments (*e.g.*, ruminants, lakes, and swamps) is characterized by clear non-equilibrium signals that yield apparent clumped isotopologue temperatures much higher than environmental methane generation temperatures (Stolper et al., 2015; Wang et al., 2015; Douglas et al., 2016; Young et al., 2017).

The abundance of the clumped isotopologue $^{13}\text{CH}_3\text{D}$ is reported as $\Delta^{13}\text{CH}_3\text{D}$, which represents the deviation (excess) of the abundance of $^{13}\text{CH}_3\text{D}$ from a stochastic distribution (*i.e.*, one in which all carbon and hydrogen isotopes are randomly distributed amongst the isotopologues $^{12}\text{CH}_4$, $^{13}\text{CH}_4$, $^{12}\text{CH}_3\text{D}$, and $^{13}\text{CH}_3\text{D}$) (Ono et al., 2014):

$$\Delta^{13}\text{CH}_3\text{D} = \frac{[^{13}\text{CH}_3\text{D}][^{12}\text{CH}_4]}{[^{13}\text{CH}_4][^{12}\text{CH}_3\text{D}]} - 1 \quad (1.5)$$

Methane isotopologues may thus provide additional information about the formation of the C–H bond, rather than just the origin of carbon or hydrogen, thereby providing information about the mechanism of methane molecule formation. Measurements of the clumped isotopologue of methane, $^{13}\text{CH}_3\text{D}$ and $^{13}\text{CH}_2\text{D}_2$, have been used to better understand the source of methane in a variety of environments (Gruen et al., 2014; Ono et al., 2014; Wang et al., 2015; Stolper et al., 2015, 2014; Eiler et al., 2014; Lopes et al., 2016; Whitehill et al., 2017; Young et al., 2016; Wang et al., 2018). However, life operates outside of equilibrium. Microbes can produce nonequilibrium clumped methane signals in laboratory studies and in the environment (Wang et al., 2015; Stolper et al., 2015). Stable isotopes of carbon and hydrogen are some of our best tools for making biogeochemical inferences, but the mechanisms that drive these nonequilibrium clumped isotope signals remain poorly understood.

1.3.2 Molecular Approaches to Studying Microbial Biogeochemistry

Molecular Clock Dating

Genetic sequence information can complement the geochemical information gleaned from the geologic past. Figure 1-2 illustrates some phylogenetics nomenclature that will be discussed in this thesis. Molecular clocks, phylogenetic trees temporally calibrated with fossil evidence, have been used to time the evolution of different plants and animals (Peterson et al., 2004; Donoghue and Benton, 2007). A phylogenetic tree provides relative divergence distances, and a molecular clock estimates the rate of evolution. Absolute timing of the molecular clock can be calibrated using independently known dates of characteristic fossils, similar to how index fossils are used to date stratigraphic columns in geologic settings (Hedges and Kumar, 2003; Donoghue and Benton, 2007; O'Reilly et al., 2015). However, microbial life does not leave behind fossils with which to date divergences. Consequently, new molecular

clock techniques have been developed that use extant gene and protein sequences to estimate divergence times of microbial lineages (Donoghue and Benton, 2007; Knoll, 2017; Parfrey et al., 2011; Schirmer et al., 2015).

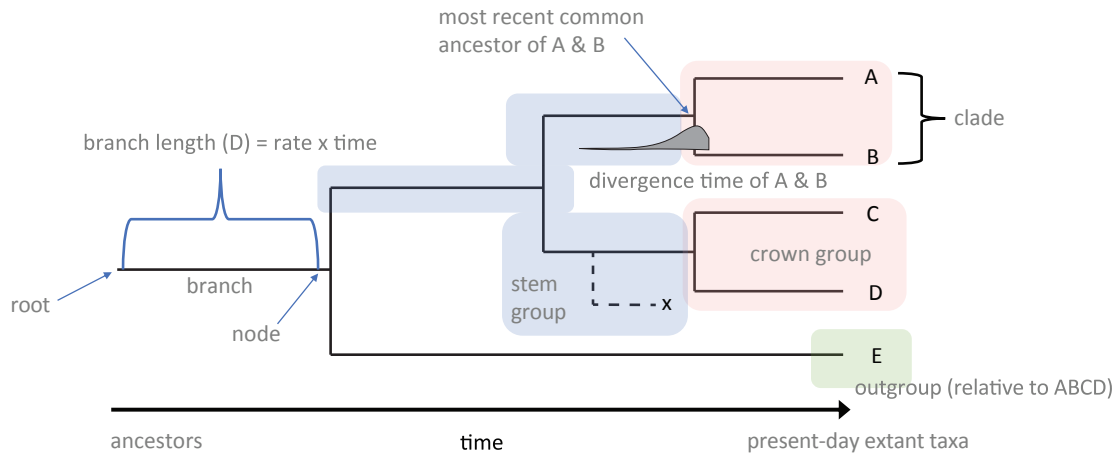


Figure 1-2: Molecular evolution nomenclature. The time axis is implicit because evolutionary changes happen between ancestor and descendants, but this is not always a real axis in that it isn't necessarily calibrated by the fossil record. The grey lines represent the branches of a tree, across which evolutionary changes occur, and the length corresponds to rate x time. Each juncture is a node. The crown group taxa are labeled relative to the stem group taxa, including extinct members. E, in green, is the outgroup of all of those taxa which we use to polarize characters.

Horizontal Gene Transfer

Horizontal gene transfer (HGT), coupled with new molecular approaches, has been proposed as one way to integrate the timing of events in microbial evolution with stronger confidence and higher precision (Wolfe and Fournier, 2018; Magnabosco et al., 2018; Dos Reis, 2018; Fournier et al., 2015). HGT has played an important role in the evolutionary history of many metabolic pathways, including methanogenesis (Fournier et al., 2009). While vertical inheritance passes genetic information from parent to offspring, HGT passes information between lineages, across any evolutionary distance. The importance of HGT in microbial evolution has become increasingly appreciated (Gogarten, 1995; Gogarten and Townsend, 2005; McInerney et al., 2008). HGT events are detected by comparing phylogenies inferred from highly conserved protein sequences (which are likely to reflect a species tree) to those of individual genes. Topological conflicts (differences in the structure of trees) between gene and species trees are candidate HGT events (Soucy et al., 2015) (Figure 1-3).

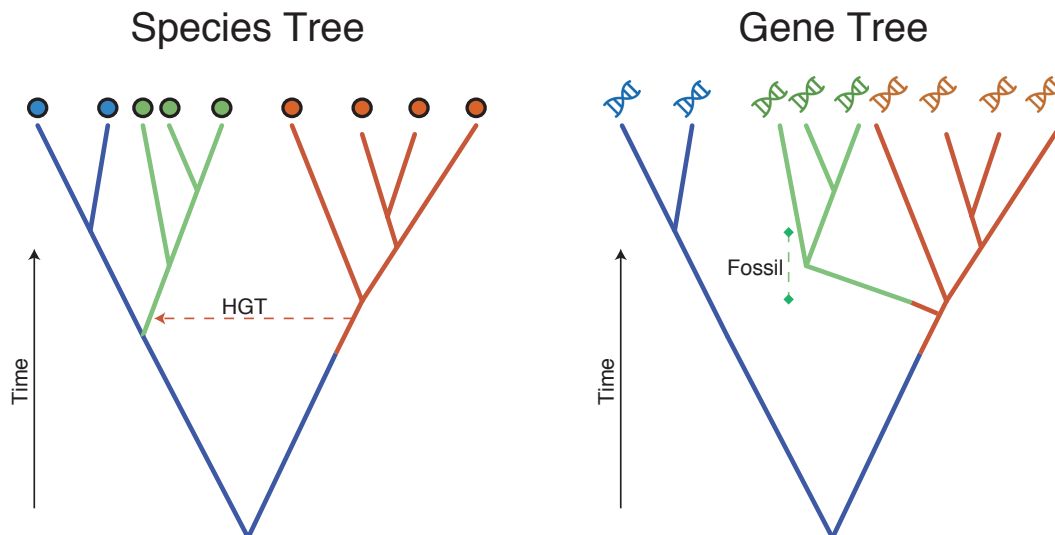


Figure 1-3: Topological conflicts between species and gene trees are used to infer HGT. Modified from (Dos Reis, 2018).

A variety of computational tools exist to detect and evaluate these events (Fournier et al., 2009, 2015; Andam et al., 2010; Zhu et al., 2014). HGT events can also be used to improve divergence time estimates, because they can be used as a "temporal scaffold" to tether other fossil or geochemical calibrations (Gold et al., 2017; Magnabosco et al., 2018; Wolfe and Fournier, 2018; Dos Reis, 2018). Figure 1-4 illustrates the pipelines used to build gene and species trees and apply molecular clock models.

1.4 Thesis Overview

The chapters that follow seek to illuminate signals of microbial metabolism in modern and ancient ecosystems. In Chapter 2 (and Appendices A, B, C), I use lab-based cultures, field sampling, and modeling to investigate possible mechanisms governing kinetic clumped isotopologue signals in microbial methane. I show that these kinetic isotope signals may not necessarily be specific to the substrate used, but do distinguish microbial methane. In Chapter 3, I explore methanogenic substrate utilization in the environment and through time by assessing the phylogenetic history of key genes in the methylotrophic methanogenic pathway. Further, I present opportunities for exploring this poorly-understood pathway in future environments and experiments. Following up on some of the challenges presented by Chapter 3, in Chapter 4 I report a novel HGT present in many microbial lineages and explain how we can leverage this event to understand the emergence of several bacterial lineages. This also has implications for our understanding of the emergence and distribution of major groups of organisms (fungi, arthropods, and bacteria) in the Proterozoic and Phanerozoic

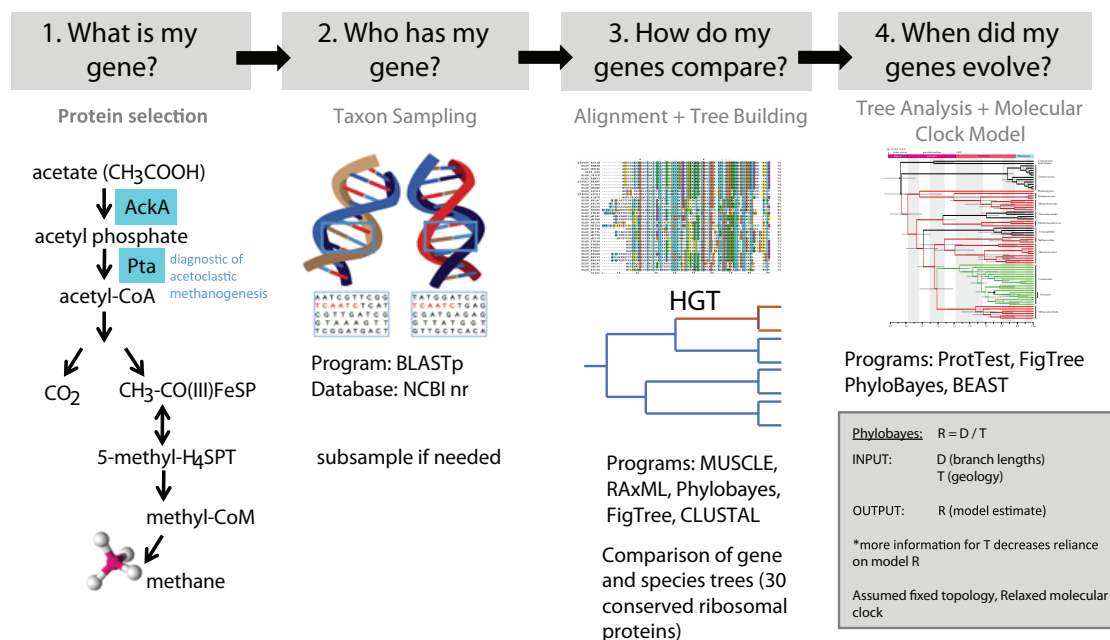


Figure 1-4: Pipeline for phylogenetic and molecular clock methods.

Eons.

In summary, this work uses both isotopic information from the modern environment and molecular evidence from genetic sequence databases to inform a better understanding of the signals of and evolutionary events involving carbon metabolism during the Proterozoic and Phanerozoic Eons. Clumped isotopologues add dimensionality to conventional carbon and hydrogen measurements in explaining the origin of methane produced by microbes in laboratory and natural systems. Substrate-specific methanogens may have diversified, in part, due to their interactions with bacterial neighbors in ancient marine sediments. The distribution and formation of some of the simplest biomolecules (*e.g.*, methane and chitin), informed by novel and complementary geochemical and phylogenetic tools, may provide important insight into how Earth's carbon cycle evolved over geologic timescales. Moreover, the methods developed in this work are broadly applicable to further understanding the coevolution of Earth and life with respect to other biogeochemical cycles of interest.

CHAPTER 2

EXPERIMENTAL INVESTIGATION ON THE CONTROLS OF CLUMPED ISOTOPOLOGUE AND HYDROGEN ISOTOPE RATIOS IN MICROBIAL METHANE

This chapter was originally published as Gruen, D.S., Wang, D.T., Könneke, M., Topçuoğlu, B., Stewart, L., Goldhammer, T., Holden, J.F., Hinrichs, K.U., Ono, S. (2018). Experimental investigation on the controls of clumped isotopologue and hydrogen isotope ratios in microbial methane. *Geochimica et Cosmochimica Acta* 237: 339–356.

D.S.G, M.K., K.U.H., L.C.S., B.T., J.F.H. designed and/or conducted the culturing experiments, D.S.G. and D.T.W. performed isotopic analyses, D.S.G. collected and analyzed the data; D.S.G. and S.O. performed modeling and wrote the manuscript with input from all authors.

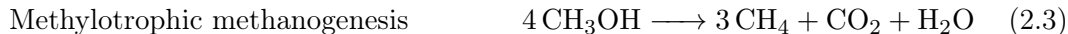
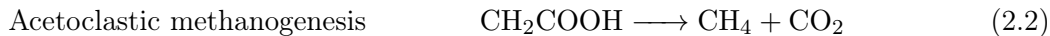
2.1 Abstract

The abundance of methane isotopologues with two rare isotopes (*e.g.*, $^{13}\text{CH}_3\text{D}$) has been proposed as a tool to estimate the temperature at which methane is formed or thermally equilibrated. It has been shown, however, that microbial methane from surface environments and from laboratory cultures is characterized by low $^{13}\text{CH}_3\text{D}$ abundance, corresponding to anomalously high apparent $^{13}\text{CH}_3\text{D}$ equilibrium temperatures. We carried out a series of batch culture experiments to investigate the origin of the non-equilibrium signals in microbial methane by exploring a range of metabolic pathways, growth temperatures, and hydrogen isotope compositions of the media. We found that thermophilic methanogens (*Methanocaldococcus jannaschii*, *Methanothermococcus thermolithotrophicus*, and *Methanocaldococcus bathoardescens*) grown on $\text{H}_2 + \text{CO}_2$ at temperatures between 60 and 80°C produced methane with $\Delta^{13}\text{CH}_3\text{D}$ values (defined as the deviation from stochastic abundance) of 0.5-2.5‰, corresponding to apparent $^{13}\text{CH}_3\text{D}$ equilibrium temperatures of 200-600°C. Mesophilic methanogens (*Methanosarcina barkeri* and *Methanosarcina mazei*) grown on $\text{H}_2 + \text{CO}_2$, acetate, or methanol produced methane with consistently low $\Delta^{13}\text{CH}_3\text{D}$ values, down to -5.2‰. Closed system effects can explain part of the non-equilibrium signals for methane from thermophilic methanogens. Experiments with *M. barkeri* using D-spiked water or D-labeled acetate (CD_3COO^-) indicate that 1.6-1.9 out of four H atoms in methane originate from water, but $\Delta^{13}\text{CH}_3\text{D}$ values of product methane only weakly correlate with the D/H ratio of medium water. Our experimental results demonstrate that low $\Delta^{13}\text{CH}_3\text{D}$ values are not specific to the metabolic pathways of methanogenesis, suggesting that they could be produced during enzymatic reactions common in the three methanogenic pathways, such as the reduction of methyl-coenzyme M. Nonetheless C–H bonds inherited from precursor methyl groups may also carry part of non-equilibrium signals.

2.2 Introduction

Methane is significant to the global carbon cycle (*e.g.*, Alperin and Blair, 1992), a potent greenhouse gas (*e.g.*, Wecht et al., 2014), a source of energy (*e.g.*, Whiticar, 1990), and a potential biosignature both for the deep biosphere (*e.g.*, Inagaki et al., 2015) and planetary missions (*e.g.*, Webster et al., 2015). The greatest natural source of methane to the atmosphere is produced by microbes in anoxic environments such as swamps, sediments, rice paddies, and ruminant tracts (Klapp et al., 2010). Microbial methanogenesis also contributes the majority of methane to oceanic gas hydrates, the largest reservoir of methane on Earth (Kvenvolden, 1993; Thauer et al., 2008).

Three major pathways are known for microbial methanogenesis (Thauer, 1998):



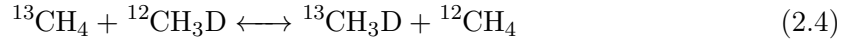
Hydrogenotrophic methanogenesis (2.1) is the reduction of CO_2 with H_2 as an electron donor, which is employed by at least six of the seven known orders of methanogens (Baptiste et al., 2005). Although not as common, methanogens capable of hydrogenotrophic methanogenesis can use formate as both a source of carbon and as an electron donor (Baptiste et al., 2005). Acetoclastic methanogenesis (2.2), the disproportionation of acetate to CH_4 and CO_2 , is restricted to the genera *Methanosarcina* and *Methanosaeta* within the order Methanosarcinales. Hydrogenotrophic and acetoclastic methanogenesis are the two most common forms of microbial methanogenesis today (Conrad, 2005). Methyлотrophic methanogenesis (2.3), used by Methanosarcinales and Methanomassiliicoccales, uses methyl compounds (*e.g.*, methanol, methyl amines, dimethylsulfide, or methylthiols) as substrates (Baptiste et al., 2005; Penger et al., 2012). Nonetheless, all methanogenic archaea possess the enzyme methyl-coenzyme M reductase (MCR) which catalyzes the final step of methanogenesis (*e.g.*, Ermler et al., 1997; Grabarse et al., 2000; Scheller et al., 2013; Wagner et al., 2016).

Carbon ($^{13}\text{C}/^{12}\text{C}$) and hydrogen (D/H) isotope ratios have been widely used to identify the origin of methane in the environment (Blair and Carter, 1992; Whiticar, 1999; Conrad et al., 2009; McCalley et al., 2014; Blaser and Conrad, 2016). However, it is often challenging to accurately determine the methane source since the isotopic composition of methane depends upon carbon and hydrogen sources as well as isotope fractionation during formation processes (Waldron et al., 1999; Valentine et al., 2004; Yoshioka et al., 2008; Kawagucci et al., 2014).

Previous culture experiments using D-labeled water (Daniels et al., 1980) or D-spiked water (Kawagucci et al., 2014) indicate that hydrogen in hydrogenotrophic methane is primarily derived from the hydrogen in water with only minor contribution from hydrogen in hydrogen gas (H_2). The contribution from H_2 can be explained by the production of metabolic water (Sugimoto and Wada, 1995) since the production of one mole of methane yields two moles of water ($\text{CO}_2 + 4 \text{H}_2 \longrightarrow \text{CH}_4 + 2 \text{H}_2\text{O}$), and the high specific rate of methanogenesis results in rapid turnover of intracellular water. The residence time of intracellular water is estimated to be as short as a few seconds (much lower than the doubling time of cells) during exponential growth (Kawagucci et al., 2014). The δD value of intracellular water can also be influenced by exchange between H_2 and H_3O^+ , which can be catalyzed by hydrogenase enzymes (Burke, 1993; Valentine et al., 2004). Alternatively, the direct transfer of hydrogen in H_2 into CH_4 can be mediated by the enzyme methylenetetrahy-

dromethanopterin dehydrogenase (Schworer et al., 1993; Schleucher et al., 1994; Klein et al., 1995b,a; Klein and Thauer, 1995; Hartmann et al., 1996). In contrast, experiments with acetoclastic methanogens indicate that up to three out of four hydrogen atoms in methane are derived from the methyl group of acetate, as implied in the stoichiometry of Reaction (2.2) (Pine and Barker, 1956).

In addition to the ratios of $^{13}\text{C}/^{12}\text{C}$ and D/H of methane, measurements of the doubly isotope substituted isotopologue, $^{13}\text{CH}_3\text{D}$ and/or $^{12}\text{CH}_2\text{D}_2$, have recently been applied as tools to constrain the source of methane in a variety of environments (Stolper et al., 2013, 2014; Wang et al., 2015; Douglas et al., 2016; Wang et al., 2016; Young et al., 2016; Douglas et al., 2016; Whitehill et al., 2017; Young et al., 2017). The following reaction shows the equilibrium among four methane isotopologues, including $^{13}\text{CH}_3\text{D}$:



Here, its equilibrium constant (K_T) can be written as:

$$K_T = \frac{[^{13}\text{CH}_3\text{D}][^{12}\text{CH}_4]}{[^{12}\text{CH}_3\text{D}][^{13}\text{CH}_4]} \quad (2.5)$$

The value of K_T primarily depends on temperature, and it approaches unity at high temperatures (1.0002 at 1,000°C), but is about 1.0057 at 25°C as estimated by molecular simulations (Ma et al., 2008; Stolper et al., 2015; Wang et al., 2015; Liu and Liu, 2016). Thus, the precise measurements of four isotopologues' abundance ($^{12}\text{CH}_4$, $^{13}\text{CH}_4$, $^{12}\text{CH}_3\text{D}$, and $^{13}\text{CH}_3\text{D}$) were thought to provide an estimate of the temperature at which the methane gas was formed or thermally equilibrated. Initial studies using high-resolution mass-spectrometry demonstrated that this new isotopologue thermometer provides a range of temperatures that are consistent with formation temperatures for methane samples from geologic environments, such as natural gas reservoirs (Stolper et al., 2014). Later studies, however, showed that methane sampled from surface environments (*e.g.*, ruminants, lakes, and swamps) is characterized by clear non-equilibrium signals that yield apparent clumped isotopologue temperatures higher than environmental methane generation temperatures (Stolper et al., 2015; Wang et al., 2015; Douglas et al., 2016; Young et al., 2017).

These studies also showed that the degree of methane isotopologue disequilibrium is correlated with D/H-isotope disequilibrium between H_2O and CH_4 (*i.e.*, $\text{CH}_3\text{D} + \text{H}_2\text{O} \longleftrightarrow \text{CH}_4 + \text{HDO}$). To explain this observed relationship, Wang et al. (2015) and Stolper et al. (2015) both presented a mathematical model that considered metabolic reversibility, which is defined as the ratio of backward to forward fluxes through an enzymatically-mediated reaction sequence. These models were based on earlier models for sulfur isotope effects of sulfate reducers (Rees, 1973; Farquhar et al., 2007; Sim et al., 2011). By choosing the appropriate fractionation factors, these models can describe isotopologue compositions of

microbial methane between kinetic and equilibrium end-members, corresponding to unidirectional and reversible reactions, respectively. Accordingly, both studies attributed the origin of kinetic clumped isotope signals intrinsic to one or more enzymatic reactions in the methanogenic pathways. The application of transition state theory (Bigeleisen, 1949) can explain $^{13}\text{CH}_3\text{D}$ abundance between equilibrium and stochastic ($\Delta^{13}\text{CH}_3\text{D} > 0\text{‰}$), but anti-clumped $^{13}\text{CH}_3\text{D}$ abundance ($\Delta^{13}\text{CH}_3\text{D} < 0\text{‰}$) requires the mixing of methane reservoirs with an often unreasonably large range of bulk δD and $\delta^{13}\text{C}$ values or a physical mechanism, including quantum mechanical tunneling (Wang et al., 2015; Whitehill et al., 2017; Young et al., 2017) ($\Delta^{13}\text{CH}_3\text{D}$ is a measure of excess $^{13}\text{CH}_3\text{D}$ as defined later in Equation 2.8). For the doubly deuterated isotopologue CH_2D_2 , purely statistical combinational-effects can also produce large apparent depletions in CH_2D_2 (Röckmann et al., 2016; Yeung, 2016; Young et al., 2017).

Recent work cultivating methanogens produced isotopologue compositions consistently out of isotopic equilibrium (Douglas et al., 2016; Young et al., 2017). These results were in agreement with previous culture studies (Stolper et al., 2015; Wang et al., 2015), but highlighted the need for further assessment of the mechanisms that control microbial $^{13}\text{CH}_3\text{D}$ compositions. In particular, the source of H in CH_4 for acetoclastic and methylotrophic methanogenesis remained uncertain (Douglas et al., 2016, 2017). The goal of this work is to better characterize the kinetic $^{13}\text{CH}_3\text{D}$ effects that lead to these generally low $\Delta^{13}\text{CH}_3\text{D}$ values, specifically during microbial methanogenesis, using a comprehensive set of metabolic pathways and temperatures. We investigated this with a series of batch culture experiments to test the effect of 1) species (*Methanothermococcus thermolithotrophicus*, *Methanocaldococcus jannaschii*, *Methanocaldococcus bathoardescens*, *Methanosarcina barkeri*, and *Methanosarcina mazei*), 2) temperature (from 30 to 85°C), and 3) substrate ($\text{H}_2 + \text{CO}_2$, acetate, and methanol). We also investigated the effect of closed-system processes as well as D/H ratios of medium water to test if apparent high-temperature signals are produced by mixing of two or more pools of methane (or its precursors), as mixing has been shown to produce a bias in the clumped isotopologue temperature estimate (Stolper et al., 2015; Wang et al., 2015; Douglas et al., 2016).

2.3 Methods

2.3.1 Laboratory Culture Experiments

Table 2-1 summarizes all culture experiments conducted in this study as well as results from our earlier experiments presented in Wang et al. (2015). Descriptions of specific experimental conditions are provided below.

Table 2-1: Summary of experiments.

Experiment (Method)	Organism(s)	Purpose	Variables	Substrates Used	Presented in
Temperature Series*	<i>Methanocaldococcus jannaschii</i> , <i>Methanocaldococcus bathoardescens</i> , <i>Methanothermococcus thermolithotrophicus</i>	Effect of growth temperature	Growth Temperature	H ₂ + CO ₂	Figures 2-1, 2-9 Tables 2-2, 2-3
Time Series	<i>Methanocaldococcus bathoardescens</i>	Closed system isotope effects	Incubation Time	H ₂ + CO ₂	Figures 2-1, 2-2, 2-9 Tables 2-2, 2-3
Substrate Series	<i>Methanosarcina barkeri</i> , <i>Methanosarcina mazei</i>	Substrate & Pathway effects	Substrate	H ₂ + CO ₂ , acetate, methanol	Figures 2-1, 2-3, 2-4, 2-5, 2-9 Tables 2-2, 2-3
Deuterated Water	<i>Methanosarcina barkeri</i>	Hydrogen source	δ D Water	H ₂ + CO ₂ , acetate, methanol	Figures 2-1, 2-3, 2-5, 2-6, 2-8, 2-9 Tables 2-2, 2-3
Deuterated Acetate	<i>Methanosarcina barkeri</i>	Hydrogen source	CD ₃ COOD Spike	acetate	Figures 2-1, 2-4, 2-8, 2-9 Tables 2-2, 2-3

*a part of data was reported in Wang et al. (2015)

Temperature Series Experiments

Pure cultures of methanogens were grown in duplicate in batch cultures at a range of temperatures (30-85°C). Three different hydrogenotrophic methanogens were selected based on their growth kinetics and optimum growth temperatures: *Methanothermococcus thermolithotrophicus*, *Methanocaldococcus jannaschii*, and *Methanocaldococcus bathoardescens*. Cultures of *M. thermolithotrophicus* and *M. jannaschii* were purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). *M. bathoardescens* was originally isolated from vent fluid at Axial Volcano, Juan de Fuca Ridge, and maintained in culture at the University of Massachusetts, Amherst (Stewart et al., 2015).

Culture medium was prepared following the recipe for DSMZ medium 282 according to Stewart et al. (2015). The headspace was filled with H₂:CO₂ (in a ratio of 80:20 by volume) at 2 bar absolute pressure. For each experiment, 5 mL of inoculum from a culture in the exponential growth phase was added to a sample vial containing 50 mL media. *M. thermolithotrophicus* was grown at 30, 40, 50, and 60°C, *M. jannaschii* was grown at 70 and 80°C, and *M. bathoardescens* was grown at 85°C. All cultures were incubated in 140 mL rubber-stoppered glass serum vials in forced-air convection ovens. Cell concentrations were monitored by cell counts with a Petroff-Hauser counting chamber and phase-contrast light microscope to determine the growth kinetics as a function of temperature (Stewart et al., 2015). Experiments for isotope measurement were stopped at a time when stationary phase was reached (5 to 64 hours, as measured in prior studies and replicated in our laboratory Huber et al., 1982; Jones et al., 1983; Ver Eecke et al., 2013). Most CO₂ (>95%) was converted to CH₄ as indicated by gas chromatography (GC) measurements of carbon dioxide and methane in the headspace gas.

Time Series Experiments

M. bathoardescens was grown under an $\text{H}_2:\text{CO}_2$ (80:20) atmosphere in replicate batch cultures at 80°C to study the effects of growth phase and closed system on $^{13}\text{CH}_3\text{D}$ and to estimate instantaneous clumped isotopologue fractionation factors. Culture medium was prepared as above. Methane was sampled and analyzed by GC from batch cultures at time points corresponding to fractional conversion of 6, 10, 68, and 77% of the initial carbon dioxide to methane (Table 2-1). The fractional conversion was calculated by dividing the volume of methane produced at the conclusion of the experiment by the volume of methane expected.

Substrate Series Experiments

To determine the effect of different metabolic pathways, established batch culture incubations of *Methanosarcina barkeri* were grown on three different substrates: $\text{H}_2 + \text{CO}_2$, (referred to as hydrogenotrophic cultures hereafter), methanol (methylotrophic cultures), and acetate (acetoclastic cultures). Cultures of *M. barkeri* (strain DSM-800) were purchased from the DSMZ (Braunschweig, Germany). The growth medium was prepared according to the recipe for DSMZ medium 120 (Balch et al., 1979). For hydrogenotrophic cultures, the headspace was filled with 1.5 absolute bar of $\text{H}_2:\text{CO}_2$ (80:20) gas mix. For acetoclastic and methylotrophic cultures, the headspace was filled with 1.5 absolute bar of $\text{N}_2:\text{CO}_2$ (70:30) gas mix and the medium was amended with 30 mM of Na-acetate or 250 mM of methanol, respectively. Cultures were incubated in duplicate near room temperature. Two sets of experiments were carried out. The first set of experiments (Set 1) was intended to provide preliminary data, and thus temperature was not strictly controlled over the course of the experiment (cultures exposed to ambient temperatures between 21 and 38°C), and the medium contained yeast extract. Nonetheless, all bottles in this series were subjected to identical environmental conditions. A second set of cultures (Set 2) was prepared and incubated under close monitoring at constant temperature (38°C). Yeast extract (YE) and casitone were omitted from medium unless otherwise noted. At the end of the experiment, cultures were killed with 1M NaOH to prevent any additional methanogen activity.

D-label and D-spike Experiments

To constrain the source of hydrogen in the hydrogenotrophic, methylotrophic, and acetoclastic pathways, a subset of *M. barkeri* cultures was also spiked with either 15 or 30 μL of D_2O per one liter of media. Additionally, acetoclastic cultures were prepared containing 10, 50, or 100% (molar fraction) deuterated acetic acid (CD_3COOD , 99% purity, Sigma-Aldrich, St. Louis, MO).

Table 2-2: Results for methanogen culture experiments.

Methanogen	Substrate ^{*1}	T (°C)	δD_{H_2O}	$\delta^{13}C_{CH_4}$	$\pm \text{‰}$	δD_{CH_4}	$\pm \text{‰}$	$\delta^{13}C_{H_3D}$	$\pm \text{‰}$	Incuba- tion (hrs)	Methane (mL) ^{*7}
Temperature Series Experiments (TS)											
<i>M. jannaschii</i> ^{*4}	H ₂ + CO ₂	80	-49.0	-18.79	0.03	-415.46	0.05	2.29	0.23	5.25 ^{*5}	NA
<i>M. bathoardescens</i> ^{*4}	H ₂ + CO ₂	85	-49.0	-12.58	0.07	-417.80	0.07	1.03	0.45	8 ^{*5}	NA
<i>M. thermolithotrophicus</i> ^{*4}	H ₂ + CO ₂	40	-49.0	-16.47	0.04	-427.76	0.04	1.38	0.34	28 ^{*5}	NA
<i>M. thermolithotrophicus</i>	H ₂ + CO ₂	30	-49.0	-17.05	0.08	-421.44	0.12	1.56	0.28	64 ^{*5}	NA
<i>M. thermolithotrophicus</i>	H ₂ + CO ₂	60	-49.0	-17.15	0.06	-409.77	0.05	0.66	0.28	6 ^{*5}	NA
<i>M. thermolithotrophicus</i> ^{*4}	H ₂ + CO ₂	60	-49.0	-17.05	0.05	-409.84	0.05	0.54	0.28	6 ^{*5}	NA
Time Series Experiments (CS)											
<i>M. bathoardescens</i>	H ₂ + CO ₂ (6%)	80	-49.6	-17.82	0.07	-350.30	0.21	2.13	0.29	2.75	NA
<i>M. bathoardescens</i>	H ₂ + CO ₂ (10%)	80	-49.6	-18.25	0.10	-347.63	0.10	2.30	0.55	3.25	NA
<i>M. bathoardescens</i>	H ₂ + CO ₂ (68%)	80	-49.6	-3.91	0.04	-396.41	0.04	2.35	0.21	3.5	NA
<i>M. bathoardescens</i>	H ₂ + CO ₂ (77%)	80	-49.6	-3.74	0.03	-402.25	0.03	2.45	0.29	4	NA
<i>Methanosarcina</i> Substrate, Spike, and Temperature Experiments - set 1^{*3}											
<i>M. barkeri</i> ^{*4}	H ₂ + CO ₂	21-38	-51.2	-59.90	0.05	-418.40	0.05	-1.34	0.22	336 ^{*6}	NA
<i>M. barkeri</i> ^{*4}	H ₂ + CO ₂	21-38	-51.2	-59.30	0.07	-422.67	0.07	-1.08	0.63	336 ^{*6}	NA
<i>M. barkeri</i>	H ₂ + CO ₂	21-38	100 ^{*3}	-59.15	0.06	-340.47	0.05	-1.32	0.23	336 ^{*6}	NA
<i>M. barkeri</i>	H ₂ + CO ₂	21-38	260 ^{*3}	-60.93	0.11	-201.10	0.11	-2.35	0.56	336 ^{*6}	NA
<i>M. barkeri</i>	methanol	21-38	-51.2	-116.30	0.11	-372.46	0.11	-5.16	0.48	336 ^{*6}	NA
<i>M. barkeri</i>	acetate	21-38	-51.2	-66.83	0.08	-317.08	0.09	-2.87	0.42	336 ^{*6}	NA
<i>M. barkeri</i>	acetate	21-38	-51.2	-66.78	0.34	-313.61	0.55	-2.97	1.49	336 ^{*6}	NA
<i>Methanosarcina</i> Substrate, Spike, and Temperature Experiments - set 2											
<i>M. barkeri</i>	H ₂ + CO ₂	38	85.4	-62.02	0.05	-373.86	0.05	-3.06	0.28	730	22
<i>M. barkeri</i>	H ₂ + CO ₂	38	322.1	-58.08	0.05	-239.13	0.05	-2.91	0.23	730	39
<i>M. barkeri</i>	H ₂ + CO ₂	38	-50.5	-57.40	0.05	-452.14	0.05	-3.82	0.32	730	30
<i>M. barkeri</i> ^{*2}	H ₂ + CO ₂	21	-49.4	-70.52	0.05	-425.43	0.05	-1.85	0.23	730	15
<i>M. barkeri</i>	methanol	38	-39.1	-117.01	0.05	-370.16	0.05	-4.92	0.40	72	22
<i>M. barkeri</i>	methanol	38	196.7	-116.32	0.06	-333.09	0.06	-4.82	0.27	72	33
<i>M. barkeri</i>	methanol	38	334.7	-118.10	0.05	-273.19	0.06	-4.17	0.36	72	96
<i>M. barkeri</i>	acetate	38	105.1	-73.74	0.07	-300.93	0.07	-2.75	0.47	730	11
<i>M. barkeri</i>	acetate	38	359.7	-67.68	0.05	-259.54	0.05	-2.43	0.22	730	9
<i>M. barkeri</i>	acetate (+YE)	38	-35	-72.86	0.10	-343.12	0.10	-3.10	0.45	730	12
<i>M. mazei</i>	H ₂ + CO ₂	38	-51.1	-56.07	0.05	-460.03	0.05	-3.46	0.22	730	23
<i>M. mazei</i>	acetate	38	-34.1	-60.40	0.05	-314.90	0.05	-1.70	0.20	730	11
<i>M. mazei</i>	methanol	38	-23	-120.82	0.04	-363.34	0.04	-4.47	0.20	72	86

^{*1} Substrate (%) refers to the percent of substrate consumed or reaction completion.

^{*2} The culture was grown at 21°C and excluded from Figure 2-3, 2-4 and analysis.

^{*3} δD_{H_2O} values represent estimates. These were measured for Set 2.

^{*4} From Wang et al. (2015)

^{*5} Values represent an estimate based on previous culture data using this this culture strain grown at similar temperatures.

^{*6} Values represent estimated incubation time.

^{*7} Volume of methane in culture headspace (STP) at the end of the incubation time. NA (not analyzed) is recorded for all cultures that were not sacrificed (injected with NaOH) prior to isotope measurements.

2.3.2 Sample preparation and Isotopologue Measurements

At the completion of an experiment, 1 M NaOH was injected in each culture bottle (at a ratio of 1 mL per 10 mL of medium) to sacrifice the culture and to draw down CO₂ pressures in the headspace. Methane samples from culture experiments were measured within one year of the completion of culture experiments. Repeated measurements of NaOH-treated samples did not show measureable changes in isotope or isotopologue ratios during storage. The headspace was sampled by flushing with helium via two needles. Methane gas was purified from culture gas mixtures (mostly methane, hydrogen, and nitrogen) using an automated preparative gas chromatography system as previously described (Wang et al., 2015). For most analyses, approximately 10 mL STP of methane was used.

The abundance of isotopologues in methane samples was measured by a tunable infrared laser direct absorption spectrometer (TILDAS) that measures absorption in the infrared region of the electromagnetic spectrum corresponding to bending vibrations of C–H and C–D bonds (Ono et al., 2014; Wang et al., 2015). A typical measurement consists of eight to ten cycles of alternating measurements of reference and sample methane. Measured isotopologue ratios were averaged and 95% confidence intervals were calculated according to Student’s *t*-distribution as previously described (Wang et al., 2015).

The hydrogen isotope composition of culture medium water ($\delta\text{D}_{\text{H}_2\text{O}}$), except for that of *Methanosarcina* Set 2 experiments, was measured using a cavity ring-down spectrometer (CRDS, Picarro Inc., Santa Clara, California, USA) at the University of Massachusetts, Amherst. The $\delta\text{D}_{\text{H}_2\text{O}}$ values of the *Methanosarcina* Set 2 cultures were measured at the University of Bremen also by CRDS (Picarro L2130-i Analyzer, Picarro Inc., Santa Clara, CA, USA). The hydrogen isotope composition of H₂ was not measured. Bulk $\delta^{13}\text{C}$ of methanol and acetate were measured via LC-IRMS at the University of Bremen (Heuer et al., 2006). CO₂ in the N₂:CO₂ and H₂:CO₂ gas mixes, bicarbonate solution, and culture media were measured via isotope ratio infrared spectrometry (IRIS) at the University of Bremen. The D/H ratio of sodium acetate (CH₃COONa) was measured by high temperature conversion elemental analyzer interfaced with isotope ratio mass-spectrometer (IRMS) at University of Chicago. Typical uncertainties were 0.2 to 0.4‰ and 2 to 5‰ for $\delta^{13}\text{C}$ and δD , respectively.

2.3.3 Isotope Notation and Calibration

In this work, stable isotopic ratios of carbon and hydrogen are reported in conventional δ notation, defined as:

$$\delta^{13}\text{C} = \frac{(^{13}\text{C}/^{12}\text{C})_{\text{sample}}}{(^{13}\text{C}/^{12}\text{C})_{\text{PDB}}} - 1 \quad (2.6)$$

$$\delta\text{D} = \frac{(\text{D}/\text{H})_{\text{sample}}}{(\text{D}/\text{H})_{\text{SMOW}}} - 1 \quad (2.7)$$

where, PDB and SMOW are Pee Dee Belemnite and Standard Mean Ocean Water, respectively. The factor of 1000, which commonly appears in definitions of δ values in the geochemical literature, has been omitted from Equations 2.6 and 2.7, as it is implied by the permil (‰) symbol, in accordance with IUPAC recommendations (Coplen, 2011). Values for $\delta^{13}\text{C}$ and δD of methane analyzed via TILDAS at MIT have been calibrated against PDB and SMOW via measurements of natural gas standards NGS-1 and NGS-3 (Wang et al., 2015). Reference values for $\delta^{13}\text{C}$ and δD were taken to be -29.0‰ and -138‰ for NGS-1, and -72.8‰ and -176‰ for NGS-3, respectively (Hut, 1987).

Because the TILDAS measures ratios of methane isotopologues, bulk $\delta^{13}\text{C}$ and δD values reported in this paper are necessarily derived quantities. For samples of methane containing a mix of isotopologues at or sufficiently close to their naturally-occurring abundances, including all samples analyzed via TILDAS in this study, ratios of isotopologues are interchangeable with ratios of isotopes (*i.e.*, $^{13}\text{C}/^{12}\text{C}$ and D/H) when calculating δ values, with no difference within achievable uncertainties of isotope ratio measurements: $^{13}\text{C}/^{12}\text{C} \approx [^{13}\text{CH}_4]/[^{12}\text{CH}_4]$ and $\text{D}/\text{H} \approx \frac{1}{4} [^{12}\text{CH}_3\text{D}]/[^{12}\text{CH}_4]$. Note that the symmetry factor of $\frac{1}{4}$ cancels out when δD values are calculated via Equation 2.7.

The abundance of the clumped isotopologue $^{13}\text{CH}_3\text{D}$ is reported as $\Delta^{13}\text{CH}_3\text{D}$, which represents the deviation (excess) of the abundance of $^{13}\text{CH}_3\text{D}$ from a stochastic distribution (*i.e.*, one in which all carbon and hydrogen isotopes are randomly distributed amongst the isotopologues $^{12}\text{CH}_4$, $^{13}\text{CH}_4$, $^{12}\text{CH}_3\text{D}$, and $^{13}\text{CH}_3\text{D}$) (Ono et al., 2014).

$$\Delta^{13}\text{CH}_3\text{D} = \frac{[^{13}\text{CH}_3\text{D}][^{12}\text{CH}_4]}{[^{13}\text{CH}_4][^{12}\text{CH}_3\text{D}]} - 1 \cong \ln \frac{^{13}\text{CH}_3\text{D}}{^{13}\text{CH}_4} - \ln \frac{^{13}\text{CH}_4}{^{12}\text{CH}_4} - \ln \frac{^{12}\text{CH}_4\text{D}}{^{12}\text{CH}_4} \quad (2.8)$$

We used the following equation to derive apparent $\Delta^{13}\text{CH}_3\text{D}$ temperatures:

$$\Delta^{13}\text{CH}_3\text{D}(T) = -0.1101 \left(\frac{1000}{T} \right)^3 + 1.0415 \left(\frac{1000}{T} \right)^2 - 0.5223 \left(\frac{1000}{T} \right) \quad (2.9)$$

where T is in Kelvin. Density function theory (B3LYP) with 6-31G(d) basis set was used to estimate harmonic vibrational frequencies, and isotope fractionation factors were calculated following conventional theory by Urey (1947). A sample of methane with stochastically-distributed abundances of isotopologues has a $\Delta^{13}\text{CH}_3\text{D}$ value of zero, corresponding to an apparent equilibrium temperature of infinity for Reaction 2.4. Negative $\Delta^{13}\text{CH}_3\text{D}$ values represent "anti-clumped" signals, where the abundance of $^{13}\text{CH}_3\text{D}$ is more depleted than that expected for stochastic isotopologue distribution.

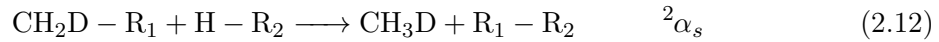
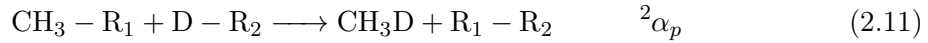
Measurements made via TILDAS give the abundances of methane isotopologues relative to a reference gas against which the samples are measured (here, a commercially-sourced cylinder of methane termed "AL1" was used as the reference gas). To express $\Delta^{13}\text{CH}_3\text{D}$ val-

ues of samples relative to the stochastic distribution requires determination of the $\Delta^{13}\text{CH}_3\text{D}$ value of the reference gas AL1. This was determined by heating AL1 in a flame-sealed glass tube in the presence of a platinum catalyst between 150 and 400°C for several days to months (Ono et al., 2014; Wang et al., 2015).

The fractionation factor (α) quantifies the difference in the relative abundance of isotopes between the substrate and the instantaneous product of a reaction. For the reduction of carbon dioxide to methane, the fractionation factor is defined as:

$$^{13}\alpha_{\text{CH}_4/\text{CO}_2} = \frac{\delta^{13}\text{C}_{\text{CH}_4} + 1}{\delta^{13}\text{C}_{\text{CO}_2} + 1} \quad (2.10)$$

Two modes of D/H fractionation characterize each hydrogen addition step during the biosynthesis of methane. For example, for addition of H (or D) onto a methyl group:



Reaction 2.11 is accompanied by a *primary* D/H isotope effect (characterized by the fractionation factor ${}^2\alpha_p$), where a D is substituted for H in the bond formed (or broken). Reaction 2.12 is accompanied by a *secondary* D/H isotope effect (with fractionation factor ${}^2\alpha_s$), where the substitution of D for H occurs on the site *adjacent* to the C–H bond being formed (or broken) and the C–D bond are carried from reactant to product. Primary D/H isotope effects are typically much larger compared to secondary isotope effects. For the reduction of methyl-coenzyme M to methane above, the secondary isotope effect is 0.84 and the primary isotope effect of the backward reaction is 0.41 (Scheller et al., 2013).

According to the rule of geometric mean (Bigeleisen, 1955), the fractionation factor for the clumped isotopologue $^{13}\text{CH}_3\text{D}$ is usually close, but not necessarily equal, to the product of carbon and hydrogen fractionation factors ($^{13-2}\alpha \simeq {}^{13}\alpha^2\alpha$). A significant departure from this rule has been observed for some *in vitro* enzyme assay experiments for doubly deuterated substrates, and attributed to quantum mechanical tunneling (*e.g.*, Srinivasan and Fisher, 1985; Amin et al., 1988; Huskey, 2006). We represent the departure from this relationship by the γ factor, which is a metric of the kinetic clumped isotope effect (Wang et al., 2015). There are two ways by which a ^{13}C -containing methyl group can acquire an H (or D) to form $^{13}\text{CH}_3\text{D}$ (analogous to Reactions 2.5 and 2.6). Thus, there are two γ factors corresponding to primary (γ_P) and secondary D/H isotope effects (γ_S):

$$^{13-2}\alpha_p = \gamma_P {}^{13}\alpha^2\alpha_p, \text{ and} \quad (2.13)$$

Table 2-3: Isotopic compositions of substrates and medium water.

Material	$\delta^{13}\text{C}\text{‰}$	$\delta\text{D}_{\text{H}_2\text{O}}\text{‰}$	$\delta\text{D}_{\text{CH}_3}\text{‰}$
$\text{H}_2 - \text{CO}_2$ gas mix ^{*1}	-34.4		
$\text{N}_2 - \text{CO}_2$ gas mix ^{*1}	-35.8		
$\text{CH}_3\text{COONa}^1$	-40.2		-123
CH_3OH^1	-49.5		
Bremen DI Water ^{*2}		-51.2	
UMass DI Water ^{*3}		-49.6	

^{*1}: These materials were used for Set 2 of substrate and D-spike experiments.

^{*2}: Bremen DI water (pre-inoculation) was used in Set 2 of substrate and D-spike experiments.

^{*3}: UMass DI water (pre-inoculation) was used in the temperature and time-series experiments.

$$^{13-2}\alpha_s = \gamma_S^{13}\alpha^2\alpha_s \quad (2.14)$$

For bond forming reactions, product methane could become anti-clumped ($\Delta^{13}\text{CH}_3\text{D} < 0$) when the value of γ is less than unity.

2.3.4 FTIR Analysis of Methane Isotopologues

A Fourier transform infrared (FTIR) spectrometer (iS5, Thermo Scientific, Waltham, Massachusetts, USA) was used to quantify the mixing ratios of deuterated isotopologues of methane (CHD_3 , CH_2D_2 and CH_3D) and non-deuterated methane (CH_4), produced in acetoclastic cultures spiked with CD_3COOD . The FTIR spectrometer has a 0.8 cm^{-1} spectral resolution, and is equipped with a gas cell that has a path length of 10 cm, volume of 70 mL, and windows of KBr. The cell was evacuated and filled with argon three times prior to injection of the sample or standard. For each measurement, 100 μL to 1 mL standard temperature and pressure (STP) of the standard or sample (culture headspace, subsampled with a gas-tight syringe, Vici Valco, Houston, Texas, USA) was injected into the cell through a small inlet valve. Reference spectra were taken on samples of ordinary CH_4 (containing D at natural abundance) and on pure ($>98\%$ purity) deuterated isotopologues (CH_3D , CH_2D_2 , and CD_3H) purchased from Cambridge Isotope Laboratories (Cambridge, MA). The mixing ratio of methane isotopologues was determined by a least squares fit in the region of the absorption spectrum between 3200 and 2800 cm^{-1} .

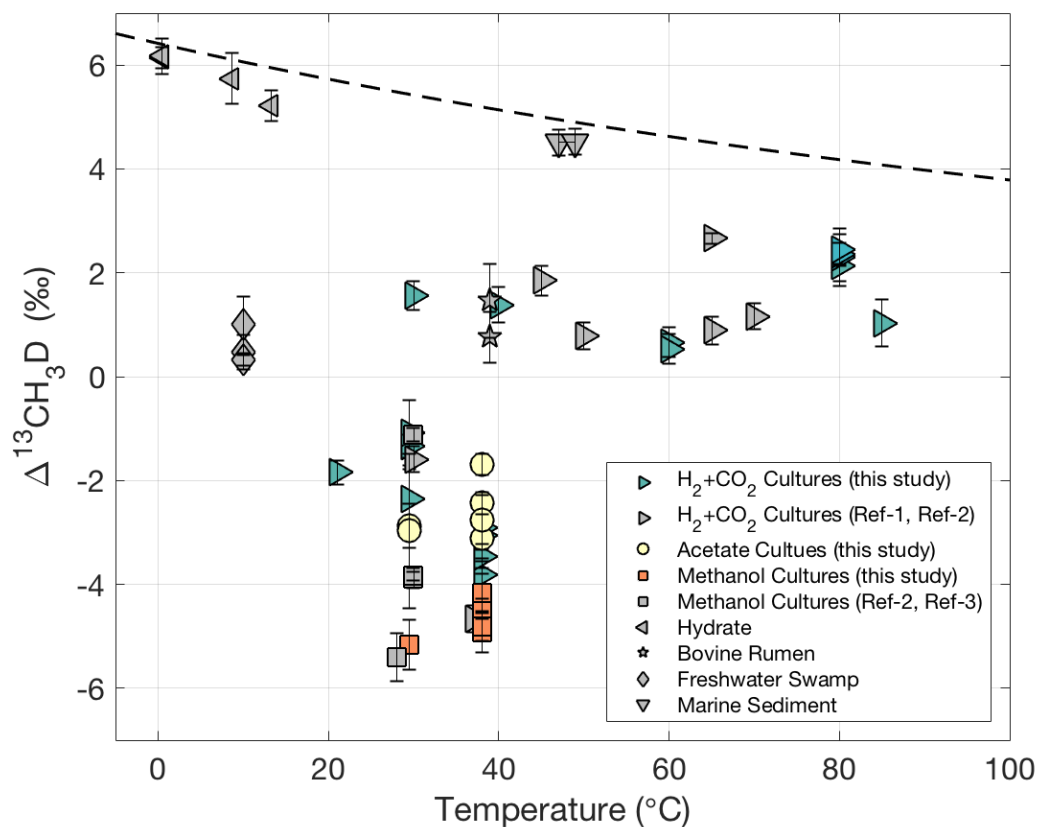


Figure 2-1: Clumped methane, $\Delta^{13}\text{CH}_3\text{D}$, plotted against environmental temperatures. The dashed line represents the equilibrium $\Delta^{13}\text{CH}_3\text{D}$ values calibrated experimentally using heated methane calibrations between 400 and 150°C (Ono et al., 2014; Wang et al., 2015) and extrapolated for lower temperatures. Colored triangles, circles, and squares represent laboratory cultures from this study. Right-facing triangles refer to $\text{H}_2 + \text{CO}_2$ cultures, circles to acetate cultures and squares to methanol cultures. A subset of samples was previously published in Wang et al. (2015) as noted in Table 2-2. For comparison to this work, previously reported $\text{H}_2 + \text{CO}_2$ and methanol cultures from Stolper et al. (2014) (Ref-1), Young et al. (2017) (Ref-2) and Douglas et al. (2016) (Ref-3) are plotted with grey symbols which correspond to the substrate used. Also plotted are environmental methane samples reported in Wang et al. (2015) (Grey symbols). Bovine rumen samples are published in Lopes et al. (2016). *In situ* temperatures for hydrate samples are calculated using depths and geothermal gradients listed in IODP reports (Riedel et al., 2006). All previously published culture data are reported according to their original measurement notation ($\Delta^{13}\text{CH}_3\text{D}$ or Δ^{18}).

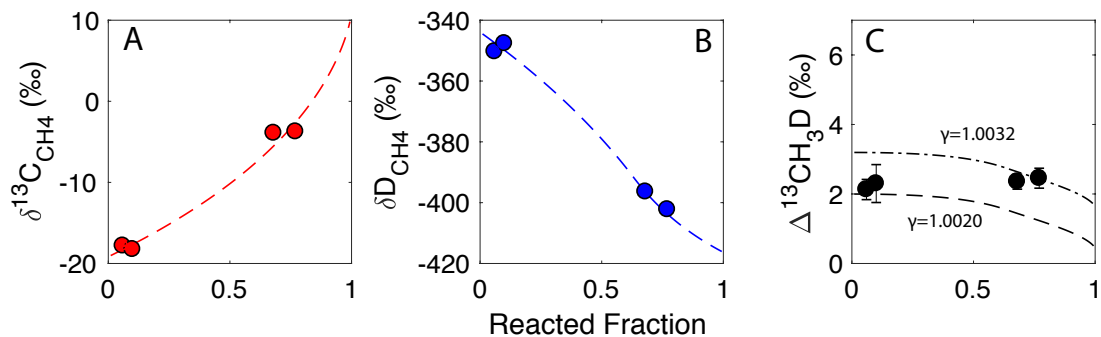


Figure 2-2: Isotope systematics of $\delta^{13}\text{C}$, δD , $\Delta^{13}\text{CH}_3\text{D}$ over the course of a batch culture experiment. Hydrogenotrophic methanogens (*Methanocaldococcus bathoardescens*) grown at 80°C. Reacted fraction refers to the fraction of carbon dioxide converted to methane. The filled circles represent methane measured from culture experiments. The dashed lines show the results of a closed system model discussed in Section 2.5.1. $\delta^{13}\text{C}$, δD and $\Delta^{13}\text{CH}_3\text{D}$ of methane are shown in A, B, and C, respectively. Results for two γ values are shown in C. See text for other fractionation factors. The $\delta^{13}\text{C}$ value of initial CO_2 was fitted to 10.9‰.

2.4 Results

Table 2-1 (Section 2.3) summarizes the tables and figures in which results for each set of methanogen culture experiments are displayed, Table 2-2 summarizes all data used in the figures that follow, and Table 2-3 summarizes the isotope composition of substrates and medium water.

2.4.1 Non-equilibrium $\Delta^{13}\text{CH}_3\text{D}$ signals of methane from microbial cultures

Microbial methane produced from pure culture experiments yielded non-equilibrium signals with corresponding apparent clumped isotope temperatures much higher than the temperatures at which the cultures were incubated (Figure 2-1). Overall, thermophilic methanogens (grown at $>40^\circ\text{C}$) produced $\Delta^{13}\text{CH}_3\text{D}$ values that are lower than those expected for equilibrium distribution (0.5 to 2.5‰), whereas mesophilic methanogens produced lower (mostly anti-clumped) signals (-5.2 to 1.6‰), consistent with limited measurements reported in Douglas et al. (2016) and Young et al. (2017).

Thermophilic methanogens (*M. jannaschii*, *M. bathoardescens*, and *M. thermolithotrophicus*) grown on $\text{H}_2 + \text{CO}_2$ between 30 and 80°C produced methane with $\Delta^{13}\text{CH}_3\text{D}$ values ranging from 2.5 to 0.5‰, corresponding to apparent clumped isotopologue temperatures of 195 to 603°C, respectively. Methane produced by *Methanosarcina* (*M. barkeri* and *M. mazei*) grown on $\text{H}_2 + \text{CO}_2$, acetate, and methanol was characterized by $\Delta^{13}\text{CH}_3\text{D}$ values ranging from -5.2 to -1.1‰, which are lower than those of methane produced by thermophilic

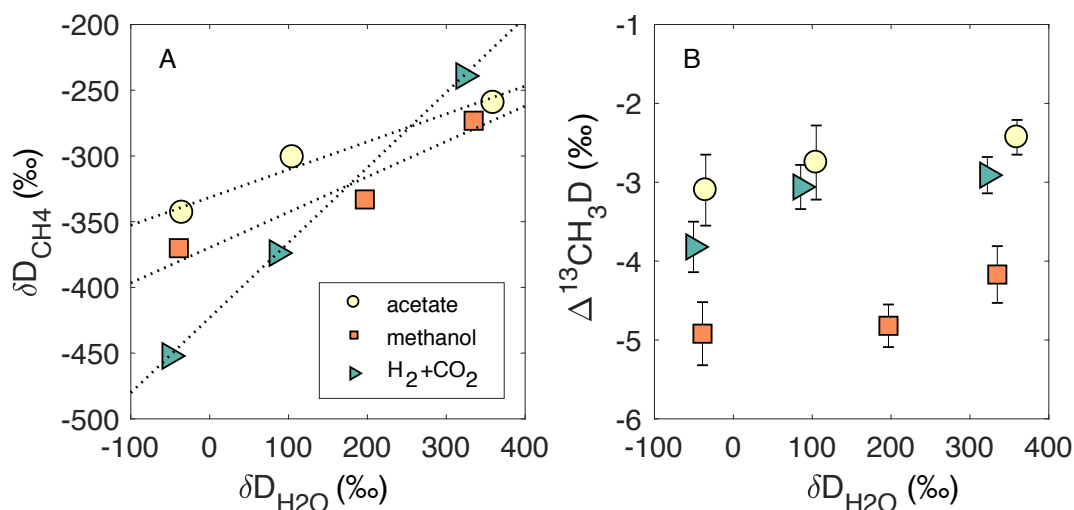


Figure 2-3: The δD and $\Delta^{13}CH_3D$ values of methane produced by *M. barkeri* in deuterium-spiked medium. Cultures were grown at 38°C, and isotopic compositions of methane are compared against δD values of media water in D-spiked experiments. Includes Set 2 data.

methanogens (Figure 2-1).

As described in the methods, Set 1 cultures were exposed to ambient temperatures between 21 and 38°C and the media contained yeast extract. Set 2 cultures were prepared and incubated under close monitoring at constant temperature (38°C). Yeast extract (YE) and casitone were omitted from media unless otherwise noted (Table 2-2), and at the end of the experiment, cultures were killed with 1M NaOH. There are some notable differences in the $\Delta^{13}CH_3D$ values between Set 1 and Set 2 experiments for cultures with *M. barkeri* and *M. mazei*. Table 2-2 shows that Set 1 exhibits slightly higher $\Delta^{13}CH_3D$ values for each substrate, most extreme for cultures grown on $H_2 + CO_2$ (as much as a 3‰ difference). *M. barkeri* cultures in Set 2 grown on $H_2 + CO_2$ at lower temperatures (21 vs. 38°C) but in the same conditions without any isotope spike exhibit higher $\Delta^{13}CH_3D$ values (by 2.0‰).

Because methanogens grown on different substrates and at different temperatures exhibit different growth rates, methane was generated more quickly or slowly for some bottles. For the Time Series experiments (TS), the culture headspace reached <1% CO_2 in as little as 5 hours for 80°C cultures and as long as 3 days for 30°C cultures. Cultures in the Time Series experiments (CS) were stopped at intervals between 2.75 and 4 hours, spanning 6-77% reaction completion (Table 2-2). Methanogens from the Substrate, Spike, and Temperature experiments (1 & 2) took much longer to generate methane. Set 2 cultures required long as two months of incubation in order to produce enough methane to be sampled and analyzed.

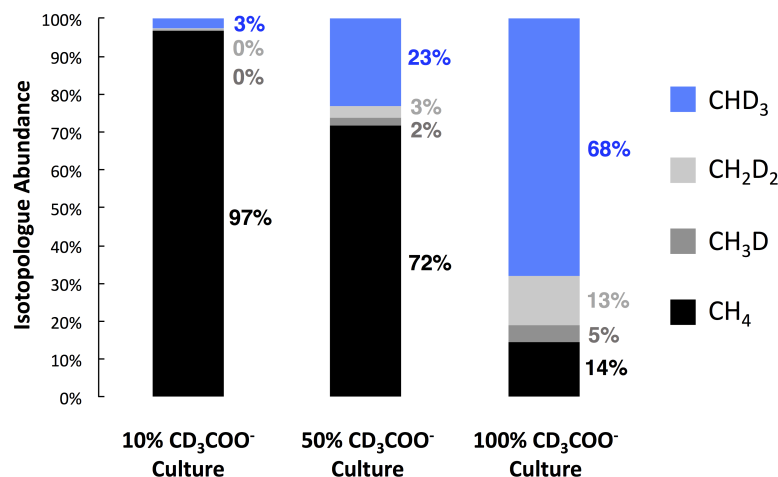


Figure 2-4: Relative isotopologue abundances of methane produced by *M. barkeri* in batch cultures spiked with CD₃COO⁻ (10, 50, 100%). Isotopologue composition was determined by FTIR, and relative abundances of methane-d_n isotopologues were calibrated against high-purity synthetic standards. CD₄ was not detected in any of the experiments.

2.4.2 Effect of a closed system on $\delta^{13}\text{C}$, $\delta\text{D}_{\text{CH}_4}$ and $\Delta^{13}\text{CH}_3\text{D}$ systematics

In order to test the potential bias in $\Delta^{13}\text{CH}_3\text{D}$ values due to closed system isotope effects, methane was sampled from batch cultures in time series experiments (Table 2-2; Figure 2-2). The $\delta^{13}\text{C}$ value of methane increased from -18.0 to -3.8‰ over the course of experiments (Figure 2-2A). This increase is consistent with closed system isotope effects. In contrast, δD values of methane decreased (-350.3 to -402.3‰) over the course of the experiment (Figure 2-2B). Our results show $\Delta^{13}\text{CH}_3\text{D}$ values remain relatively constant over the course of the experiment between 2.1 and 2.5‰ (Figure 2-2C).

2.4.3 D-spiked H₂O experiments

As the δD of water is increased by spiking the media water, the δD of product methane also increased (Figure 2-3). This illustrates the uptake of hydrogen from water to form methane, consistent with previous pure culture (Yoshioka et al., 2008; Kawagucci et al., 2014; Okumura et al., 2016) and incubation experiments (Schoell, 1980; Sugimoto and Wada, 1995). Linear regression of the data for hydrogenotrophic, acetoclastic, and methylotrophic cultures yielded the following relationships, respectively:

$$\delta D_{CH_4} = (0.571 \pm 0.011)\delta D_{H_2O} - (423.0 \pm 2.1), \quad (2.15)$$

$$\delta D_{CH_4} = (0.212 \pm 0.004)\delta D_{H_2O} - (331.5 \pm 0.8), \text{ and} \quad (2.16)$$

$$\delta D_{CH_4} = (0.269 \pm 0.005)\delta D_{H_2O} - (369.7 \pm 1.2). \quad (2.17)$$

where intercepts are in ‰. The linear fit and standard error for the slope and intercept was calculated following (York et al., 2004) by taking into account standard errors of 0.2‰ and 5‰ for δD_{CH_4} and δD_{H_2O} (assuming errors are not correlated). Hydrogenotrophic cultures yielded a higher slope (0.571) compared to methylotrophic (0.269) and acetoclastic (0.211) cultures.

The values of $\Delta^{13}CH_3D$ are weakly dependent on pathways: -2.9 to -3.8‰, -4.2 to -4.9‰, and -2.4 to -3.1‰ for hydrogenotrophic, methylotrophic, and acetoclastic cultures, respectively. For each pathway, lower $\Delta^{13}CH_3D$ values tend to be associated with lower δD_{H_2O} values (Figures 2-3 and 2-4). The linear fit and standard error for the slope and intercept was calculated as described above. Hydrogenotrophic cultures yielded a higher

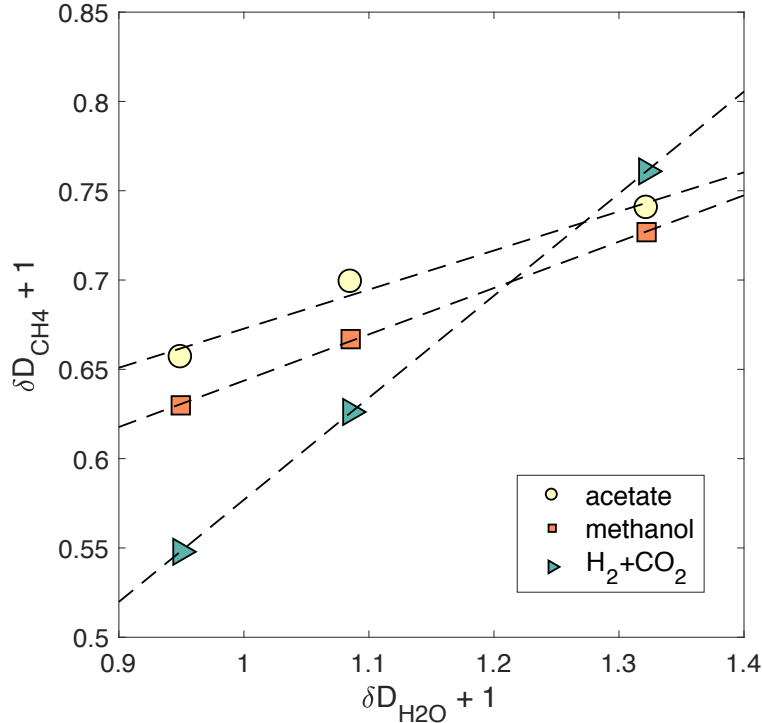


Figure 2-5: The hydrogen isotope composition of microbial methane vs. the deuterium composition of media water in spiked culture experiments. *M. barkeri* grown on three different substrates. In comparison to Figure 2-3A, the axes in this figure are plotted as $\delta D + 1$ to take into account the non-linearity in δD .

slope (0.0020 ± 0.0010) compared to methylotrophic (0.0019 ± 0.0014) and acetoclastic (0.0016 ± 0.0012) cultures.

2.4.4 D-labeled acetate experiment

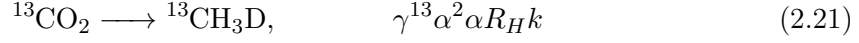
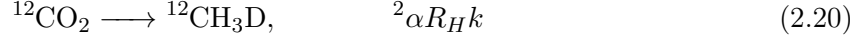
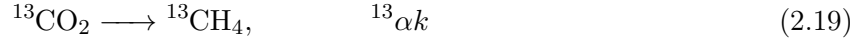
In order to track the transfer of D (or H) from the methyl group of acetate to methane, *M. barkeri* was cultured with medium spiked with CD_3COO^- . As the amount of CD_3COO^- in an acetoclastic methanogen culture increased, not only the relative abundance of CHD_3 but also that of CH_2D_2 and CH_3D increased at the expense of CH_4 (Figure 2-4). Cultures incubated with 100% CD_3COO^- produced methane comprised of a majority of triply deuterated isotopologues (68% CHD_3) but also contained CH_2D_2 , (13%) and CH_3D (5%).

2.5 Discussion

2.5.1 Closed system isotope effect does not explain non-equilibrium $\Delta^{13}\text{CH}_3\text{D}$

Methanogens in the temperature series experiment were incubated until nearly all the substrate (CO_2) had been converted to product (CH_4). As a result, the $\delta^{13}\text{C}$ value of the product methane would have increased with reaction progress, eventually reaching the $\delta^{13}\text{C}$ value of the starting CO_2 . In addition to changes in δ values due to closed system effects, it has been shown that the apparent D/H fractionation factor between methane and water changes with growth phase (Valentine et al., 2004; Kawagucci et al., 2014). This could be due to changes in the δD value of intracellular water via D/H exchange with H_2 (Burke, 1993) or the contribution from metabolic water (Kawagucci et al., 2014). Values of $\Delta^{13}\text{CH}_3\text{D}$ do track non-linearly with $\delta^{13}\text{C}$ and δD upon the mixing of two or more pools of methane (Stolper et al., 2015; Wang et al., 2015; Douglas et al., 2016), such that mixing of two methane reservoirs would result in non-equilibrium $\Delta^{13}\text{CH}_3\text{D}$ values even when $\Delta^{13}\text{CH}_3\text{D}$ values of the source reservoirs carry equilibrium signals. We sought to isolate any experimental effects introduced in the closed system and therefore tested if changing the $^{13}\text{C}/^{12}\text{C}$ and/or D/H ratios of bulk methane over the course of the reaction may also affect the $\Delta^{13}\text{CH}_3\text{D}$ value of the end product.

The production of four methane isotopologues from two isotopologues of CO_2 can be written as:



where, k is the pseudo-first order rate constant for $^{12}\text{CO}_2$ to $^{12}\text{CH}_4$, $^{13}\alpha$ and $^2\alpha$ are the carbon and hydrogen isotope fractionation factors, respectively, and R_H is the D/H ratio of source hydrogen (intracellular water, $R_H = R_{SMOW}(\delta\text{D}_{\text{H}_2\text{O}} + 1)$). For the application of equation 2.20 and 2.21, we assume that the source for H of CH_4 is intracellular water. The value of δD of intercellular water can be different from that of media due to exchange with H_2 or production of metabolic water inside the cytoplasm (Burke, 1993; Kawagucci et al., 2014). In addition, the direct transfer of H in H_2 to CH_4 was suggested (Kawagucci et al., 2014). Since the detailed mechanism of the effect of $\delta\text{D}\text{-H}_2$ is beyond the scope of this study, the above model includes the effect as the change of the $^2\alpha$ value during the course of the culture. Changing R_H would produce results identical to changing $^2\alpha$. Equations 2.18 to 2.21 were integrated numerically with three fitting parameters ($^{13}\alpha$, $^2\alpha$, and γ), and the results are shown in Figure 2-2. We used $\delta^{13}\text{C}$ and δD data to fit $^{13}\alpha$ and $^2\alpha$, and the $\Delta^{13}\text{CH}_3\text{D}$ data was used to fit γ value for the derived $^{13}\alpha$ and $^2\alpha$ values.

The best fit to the experimental data was obtained when $^{13}\alpha$ was 0.97 and $^2\alpha$ changed linearly from 0.69 to 0.57 from 0 to 25% reaction and remained a constant value of 0.57 afterwards (Figure 2-2A and B) ($\delta\text{D}_{\text{H}_2\text{O}}$ of -49.6‰ SMOW, and $\delta^{13}\text{C}_{\text{CO}_2}$ of 10.9‰). As $\delta\text{D}_{\text{H}_2\text{O}}$ does not change significantly during the course of the experiment, the change in $\delta\text{D}_{\text{CH}_4}$ cannot be explained by the closed system effect. The increasing fractionation factor at a later stage indicates $\delta\text{D}_{\text{CH}_4}$ is moving away from the value expected for equilibrium with water. Previous studies also observed similar changes in apparent D/H fractionation factors during early exponential growth phases (Valentine et al., 2004; Kawagucci et al., 2014; Okumura et al., 2016). Values of δD of methane produced during the early growth phase can be a function of $\delta\text{D}_{\text{H}_2}$ as well as $\delta\text{D}_{\text{H}_2\text{O}}$ (Kawagucci et al., 2014). For those $^{13}\alpha$ and $^2\alpha$ values, the best fit for γ was 1.0020 and 1.0032 for the early and late growth phases, respectively (Figure 2-2). The value of γ of higher than unity indicates that the rate of $^{13}\text{CH}_3\text{D}$ production is faster than the rate expected from the product of the two fractionation factors ($^{13}\alpha$, $^2\alpha$).

These derived γ values of 1.0020 and 1.0032 translate to closed-system corrected $\Delta^{13}\text{CH}_3\text{D}$ values of 2.0 and 3.2‰ respectively (corresponding to 243 and 135°C, for apparent equilibrium temperatures). These values are similar to the uncorrected values of 2.1 to 2.5‰, demonstrating that the effect of a closed system can only partially explain the non-equilibrium

$\Delta^{13}\text{CH}_3\text{D}$ signals of microbial methane. The numerical model above shows that the $\Delta^{13}\text{CH}_3\text{D}$ value of accumulated methane decreases by up to 1.2‰ over the course of reaction given constant γ values, due to the effect of mixing of methane formed during early and late exponential growth phases (Figure 2-2C).

2.5.2 Origin of H in methane from three methanogenesis pathways

The results of the D-spiked series experiments can be used to estimate the origin of C-H bonds in methane during three pathways of methanogenesis and associated deuterium isotope effects ($^2\alpha_p$ and $^2\alpha_s$). Assuming methane is formed via mixing of hydrogen atoms both from H_2O and the methyl group of acetate or methanol, the δD value of product methane can be written as:

$$\delta\text{D}_{\text{CH}_4} + 1 = ^2\alpha_p f (\delta\text{D}_{\text{H}_2\text{O}} + 1) + ^2\alpha_s (1 - f) (\delta\text{D}_{\text{CH}_3} + 1) \quad (2.22)$$

where, $^2\alpha_p$ is the kinetic fractionation factor from H_2O to CH_4 (primary D-isotope effect),

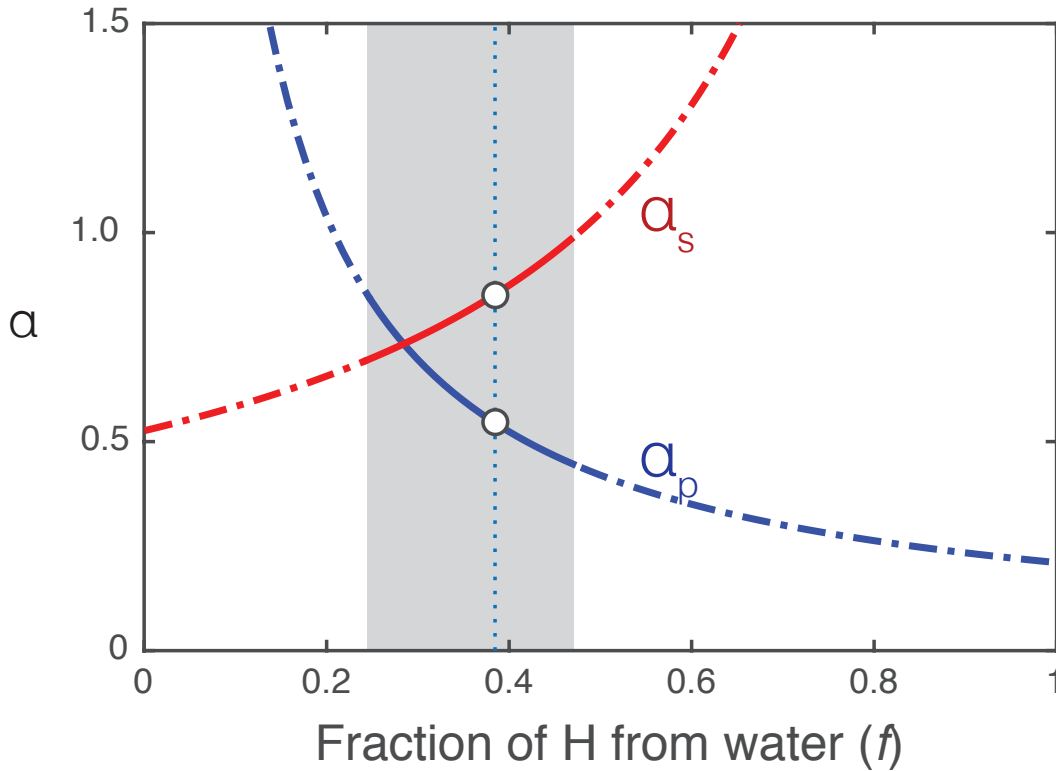


Figure 2-6: The fractionation factor (α) as a function of the fraction of hydrogen derived from water (f) in acetoclastic methanogenesis in D_2O spiked water experiments. Grey shading indicates the range of f values that satisfy γ_p and $\gamma_s < 1$ and $0.25 < f < 1$. The dashed vertical line and open circles correspond to the solution when $\gamma_s = 0.85$.

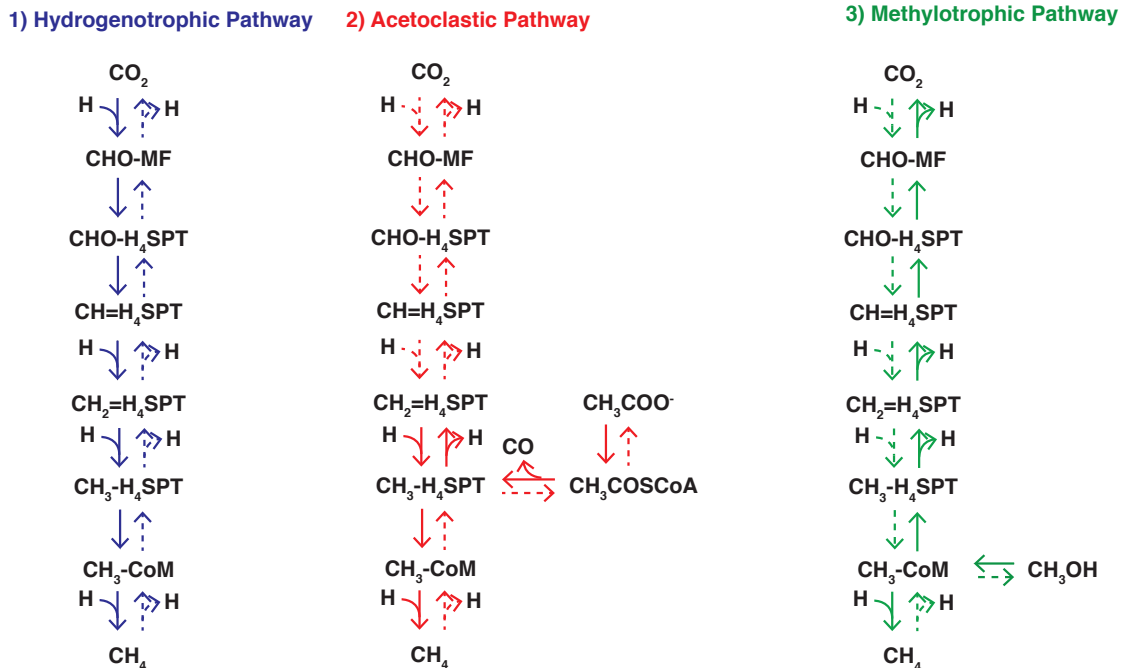


Figure 2-7: Three methanogenic pathways for *Methanosarcina*. The solid arrows represent the predominant direction of the reaction, and the dashed arrows represent the backward reaction, which is thought to be minor. Two solid arrows are used to indicate that the reactions are thought to be reversible, according to the literature, not specifically evaluated in this study. In the hydrogenotrophic pathway (blue), CO_2 is reduced by a series of four two-electron processes, each adding one H atom. C_1 compounds are carried with cofactors (MF, methanofuran, H_4SPT , tetrahydrosarcinaopterin, CoM, coenzyme M). For the acetoclastic pathway, acetate is first activated by acetyl co-A (via acetyl phosphate), acetyl co-A is split to methyl (CH_3^-) and a carbonyl moiety (CO), and the latter is oxidized to CO_2 . The methylotrophic pathway is overall a disproportionation reaction in which one methyl group is oxidized to CO_2 via a reversed methanogenic pathway, and three additional methyl groups are reduced to methane.

and $^2\alpha_s$ is the kinetic isotope fractionation factor from CH_3 (methyl-H) to CH_4 (secondary D-isotope effect) (Sessions and Hayes, 2005). The value of f is the fraction of H from H_2O . The canonical value of f is 0.25, but this can be higher when scrambling of C–H bonds occurs between C–H and intercellular H_2O .

Our experimental results yield the following relationships for $\text{CO}_2 + \text{H}_2$, acetate and methanol cultures, respectively (Figure 2-5):

$$\delta\text{D}_{\text{CH}_4} + 1 = (0.571 \pm 0.011)(\delta\text{D}_{\text{H}_2\text{O}} + 1) + (0.006 \pm 0.012), \quad (2.23)$$

$$\delta\text{D}_{\text{CH}_4} + 1 = (0.212 \pm 0.004)(\delta\text{D}_{\text{H}_2\text{O}} + 1) + (0.457 \pm 0.005), \text{ and} \quad (2.24)$$

$$\delta\text{D}_{\text{CH}_4} + 1 = (0.269 \pm 0.005)(\delta\text{D}_{\text{H}_2\text{O}} + 1) + (0.361 \pm 0.006). \quad (2.25)$$

For hydrogenotrophic methanogenesis, a small (can be zero within standard error) inter-

cept suggests all four hydrogen atoms are derived from water (*i.e.*, $f=1$), and an α_p value of 0.571 is obtained, which is within the range of previous experiments 0.55 and 0.86 (Valentine et al., 2004; Yoshioka et al., 2008; Kawagucci et al., 2014; Okumura et al., 2016).

The acetate culture (Equation 2.24) yields:

$$^2\alpha_p f = 0.211, \text{ and} \quad (2.26)$$

$$^2\alpha_s(1 - f)(\delta D_{\text{CH}_3} + 1) = 0.457. \quad (2.27)$$

The methyl-H of acetate is measured as -123‰ (*i.e.*, $\delta D_{\text{CH}_3} + 1 = 0.877$, Table 2-3). The two fractionation factors, α_p and α_s , are calculated for a given value of f in Figure 2-6. The equations (26 and 27) cannot provide a unique answer, as there are two equations with three unknowns (α_p , α_s and f). The likely range of values can be constrained, however, because isotope fractionation factors are expected to be normal (*i.e.*, $\alpha_p < 1$ and $\alpha_s < 1$), and the value of f is between $\frac{1}{4}$ and 1. This yields a range of possible values: $0.449 < ^2\alpha_p < 0.844$ and $0.695 < ^2\alpha_s < 1.0$, and $0.25 < f < 0.48$ (Figure 2-6). The secondary isotope effect for the formation of methane from methyl-coenzyme M (the last step of methanogenesis, utilized by all known methanogens Figure 2-7) is reported to be 0.84 (Scheller et al., 2013). For example, this value for $^2\alpha_s$, would yield $f=0.39$ and $^2\alpha_p=0.545$. The analysis indicates that the value of f significantly deviates from the canonical value of 0.25, which is expected from the reaction stoichiometry. Therefore, this analysis suggests that among the four hydrogen atoms in methane, up to 1.9 hydrogen atoms ($=0.48 \times 4$) are derived from water, whereas only one is required from reaction stoichiometry. Since δD_{CH_3} of methanol was not measured, we cannot carry out the same analysis for methanol experiments. The slightly higher slope for methanol cultures (0.269) compared to acetate cultures (0.212), however, suggests a similar or greater contribution of hydrogen atoms from water if we assume the same $^2\alpha_p$ value for the acetate culture.

Previous incubation studies with D-spiked water showed slopes between δD_{CH_4} and $\delta D_{\text{H}_2\text{O}}$ of 0.4 for sewage sludge (Schoell, 1980) and 0.48 to 0.61 for paddy soil (Sugimoto and Wada, 1995). These are higher than the value obtained by our experiments, indicating a higher contribution of water-H to methane in these incubation studies. Since previous experiments were enriched cultures (not pure cultures), results from previous experiments reflect a mixed contribution from hydrogenotrophic and acetoclastic methanogenesis. A previous incubation study of lake and estuary sediments using triply deuterated acetate (CD_3COO^-) showed rapid exchange of the methyl-H of acetate by methanogenic acetate metabolism (De Graaf et al., 1996). Such an exchange would contribute to the greater slope for the incubation study using natural populations, compared to pure culture experiments for this study.

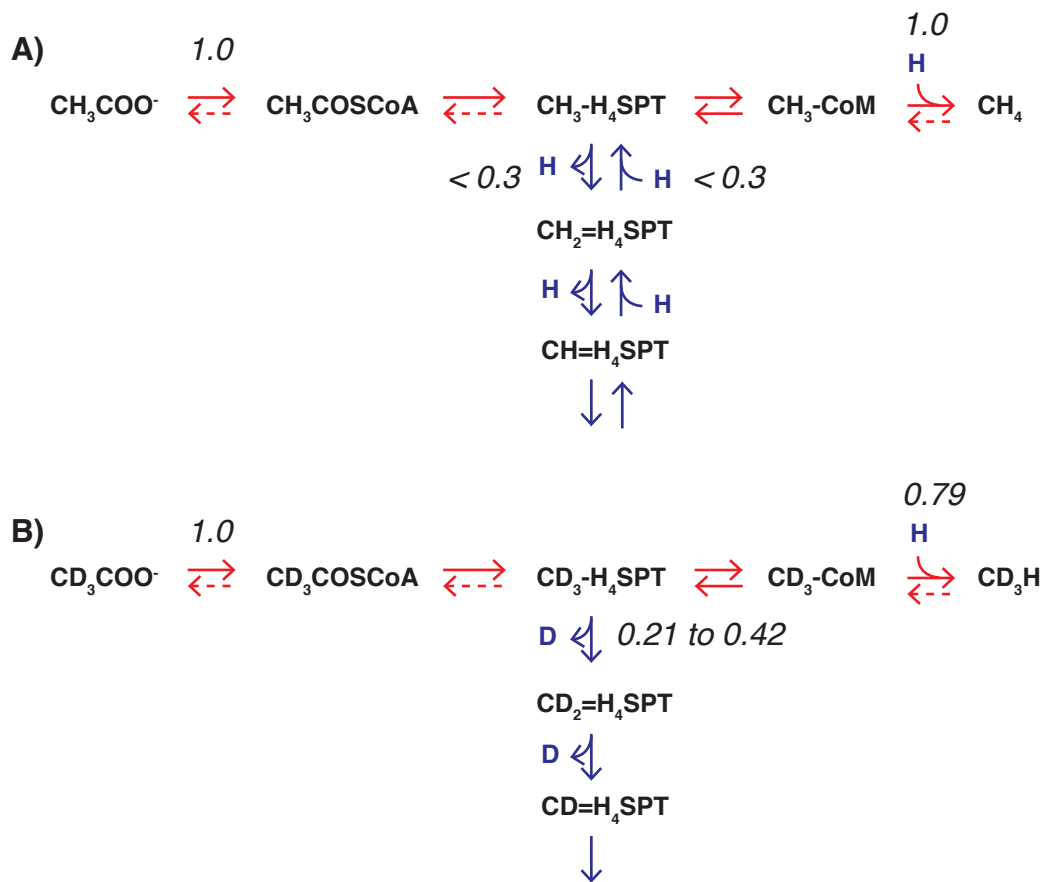


Figure 2-8: Proposed mechanism of D-isotope exchange by acetate metabolism by *M. barkeri*. The numbers in italics are relative fluxes estimated from A) D-spiked medium, and B) CD_3COO^- experiments, respectively. Reactions with two solid lines indicate that the reaction is thought to be reversible, whereas reactions with solid and dashed lines represent that the reaction proceeds predominantly the direction of the solid arrow. As in Figure 2-7, the solid arrows represent the predominant direction of the reaction, and the dashed arrows represent the backward reaction. Red lines indicate the reactions required for the acetoclastic pathway. Blue lines indicate side reactions.

2.5.3 Pathway of D-isotope exchange during acetoclastic methanogenesis

Figure 2-7 illustrates the three pathways of methanogenesis examined in this study. The solid arrows represent the predominant direction of reaction, and dashed arrows represent the backward reaction, which is thought to be minor. The two solid arrows are used where the reactions are thought to be reversible. This reversibility was inferred from our results as well as based on previous studies (*e.g.*, Thauer, 1998; Ferry, 2010). Hydrogenotrophic methanogenesis proceeds with a series of two electron reactions, which each adding one H, while the C_1 group is carried by cofactors: methanofuran (MF), tetrahydrosarcinapterin (H_4SPT), and coenzyme M (CoM). Acetoclastic methanogenesis is a disproportionation

reaction where acetate is activated to acetyl-coA (via acetyl phosphate) and split between a methyl and a carbonyl (CO) moiety, where the latter is oxidized to CO₂. The methyl group is transferred to CH₃-H₄SPT (methyl-tetrahydrosarcinaopterin), and then to CH₃-CoM (methyl-coenzyme M), and finally reduced to methane. Methylothermic methanogenesis is also a disproportionation reaction where the oxidation of one methyl group is coupled with the reduction of three methyl groups. The last step (the reduction of methyl-CoM) is common to all three pathways, and most (if not all) reactions are thought to be reversible (Thauer et al., 2008).

Our analysis of the D-spiked water experiments, together with the D-labeled acetate experiment, suggests that up to 1.9 out of four hydrogen atoms in methane are derived from water-H. In addition to one hydrogen atom that is added at the last step of methanogenesis (the reduction of methyl-CoM, Figure 2-7 and 2-8), nearly another one of three hydrogen atoms in the methyl group of acetate could be exchanged with water. Since hydrogen isotope exchange between H₂O and CH₄ is sluggish (Reeves et al., 2012), the exchange probably occurs at the CH₃ - H₄SPT (methyl-tetrahydrosarcinaopterin) intermediate via H-abstraction to form CH₂ = H₄SPT (the exchange can also occur at the methyl-CoM moiety, Scheller et al., 2013). This part of the pathway is not required for the acetoclastic metabolism but can occur as a side reaction. A similar D/H exchange mechanism was suggested to explain the observed scrambling of CD₃COO⁻ during the incubation of methanogenic sediments (De Graaf et al., 1996). Because up to 0.9 out of three methyl-H atoms are exchanged, the flux of this side reaction is estimated to be at most 0.3 H (=0.9/3) per uptake of one acetate (Figure 2-8A).

When deuterium-labeled acetate (CD₃COO⁻) was used as a substrate, the major product (68%) was the isotopologue CD₃H. However, the isotopologues CD₂H₂, CDH₃ and CH₄ were also formed (Figure 2-4). Among the four isotopologues, the fraction of CH₄ isotopologues was disproportionally high. This is because CH₄ and non-deuterated acetate are carried over from the inoculum. If all CH₄ isotopologues are from the inoculum, the proportion of CD₃H isotopologues is 79% (=68/(5+13+68)) over CD₃H, CH₂D₂ and CH₃D isotopologues. This is the maximum fraction since a small quantity (most likely <5%) could be produced from CD₃COO⁻. Following the model of D/H exchange at the CH₃-H₄SPT step as described above, this means that less than 100% yield of CD₃H is explained by the loss of CD₃ - H₄SPT to CD₂ = H₄SPT of 0.2 per one CD₃COO⁻. The reaction from CD₃-H₄SPT to CD₂ = H₄SPT and from CD₃-CoM to CD₃H would accompany deuterium isotope effects of ²α_p ²α_{s2} and ²α_{s3}, respectively. Thus, the corresponding flux for CH₃COO⁻ can be higher by the ratio of the two isotope effects (²α_s/²α_p), which could be as high as ~2 based on *in vitro* study of a similar reaction (≈0.84/0.41) (Scheller et al., 2013). Note that the medium water (and thus intercellular water) contain very little deuterium such that there is practically no back flux from CD₂ = H₄SPT to CD₃ - H₄SPT (Figure 2-8B). Thus, H/D

scrambling at the CH₃-moiety can be explained by the reverse flux of 0.2 to 0.3 based on the estimate from the D-spiked medium (0 to 0.3) and CD₃COO⁻ experiments (0.2 to 0.4) (Figure 2-8).

2.5.4 Origin of ¹³CH signals in methane from acetoclastic methanogenesis

Measured $\Delta^{13}\text{CH}_3\text{D}$ values are -1.7 to -3.1, -1.1 to -3.8, -4.2 to -5.2‰ for *Methanosarcina* cultures grown on acetate, H₂ + CO₂, and methanol, respectively, and are weakly correlated with $\delta\text{D}_{\text{H}_2\text{O}}$ of the medium (Figure 2-3B). The linear fit and standard error (95% confidence interval) was calculated following York et al. (2004). The slope ($\Delta^{13}\text{CH}_3\text{D} / \delta\text{D}$) for acetate is 0.0016 ± 0.0012 , hydrogen is 0.0020 ± 0.0010 and methanol is 0.0019 ± 0.0014 . While this is a small sample size, these are statistically not zero. These anti-clumped $\Delta^{13}\text{CH}_3\text{D}$ values can originate from 1) mixing of two or more pools of methane or its precursor with different $\delta^{13}\text{C}$ and δD values, 2) transfer of the methyl-group of acetate and methanol with pre-existing anti-clumped signals (for acetate and methanol cultures), and/or 3) intrinsic kinetic isotope effects associated with enzymatic reactions common to three pathways, such as the reduction of methyl-CoM (Figure 2-7).

It has been well known that mixing is non-linear in the clumped isotope system such that mixing of two pools of methane yields a $\Delta^{13}\text{CH}_3\text{D}$ value that is not between the two $\Delta^{13}\text{CH}_3\text{D}$ values of original two reservoirs (Eiler and Schauble, 2004; Affek and Eiler, 2006; Affek et al., 2007; Defliese and Lohmann, 2015; Röckmann et al., 2016). When two pools of methane (A and B) are mixed, the $\Delta^{13}\text{CH}_3\text{D}$ value of the mixture (δ_{mix}) can be approximated as:

$$\Delta_{mix} \approx (1 - f)\Delta_A + f\Delta_B + f(1 - f)(\delta^{13}\text{C}_A - \delta^{13}\text{C}_B)(\delta\text{D}_A - \delta\text{D}_B) \quad (2.28)$$

where f is the mixing ratio of pool B (Wang et al., 2015). The first and second terms show linear mixing between two $\Delta^{13}\text{CH}_3\text{D}$ values (Δ_A and Δ_B), whereas the third term produces a curvature following a quadratic function to f . This bias becomes the largest when the two pools of methane are mixed at a 1:1 ratio (*i.e.*, $f=0.5$), and proportional to the product of the difference of δ values between the two pools of methane. Mixing of a ¹³C- and D-enriched pool with a ¹³C- and D-depleted pool would produce a positive (low temperature) bias, whereas diagonal mixing (*e.g.*, mixing between a ¹³C-enriched and a D-depleted pool with a ¹³C-depleted and a D-enriched pool) would produce a negative (high temperature) bias in $\Delta^{13}\text{CH}_3\text{D}$. By extension, the mixing effect for doubly-deuterated clumped methane (¹²CH₂D₂) will always produce a positive (low temperature) bias (Young et al., 2016). Equation 2.28 can be used to model reaction branching (*i.e.*, producing two products) or reversible reactions (where *mix* is the source and A and B are forward and backward reactions).

A fully quantitative model for the isotope systematics for the acetoclastic pathway in Figure 2-8 is the scope of future study, but Equation 2.28 can be used to test the magnitude of any $\Delta^{13}\text{CH}_3\text{D}$ bias due to mixing. For example, mixing of two reservoirs (or fluxes) in equal portions ($f=0.5$) with $\delta^{13}\text{C}$ and δD differing by 40‰ and -300‰, respectively, can produce a non-linear bias of -3‰ ($0.5(1-0.5) \times 0.04 \times (-0.3)$). These magnitudes of kinetic ^{13}C - or D-isotope effects are possible. Therefore, an anti-clumped $\Delta^{13}\text{CH}_3\text{D}$ value of -3‰ can be produced entirely by mixing. For the acetoclastic cultures in these experiments, however, the value of $\Delta^{13}\text{CH}_3\text{D}$ was not sensitive to the δD value of the medium, which changed from -35 to 360‰ (Figure 2-3B), indicating that the mixing (between C–H bonds with water-derived H and methyl-derived H) is unlikely to be the major source of anti-clumped signals observed.

Calculations from the D-spiked cultures grown on acetate reported in this study showed that approximately up to two out of four hydrogen atoms in CH_4 are derived from water, while only one is required from stoichiometry. On average, the methyl group of acetate contributes 2 to 3 hydrogen atoms to one methane molecule, presumably also carrying its original ^{13}C –D signal. Although the degree of ^{13}C –D clumping of the methyl group of acetic acid (or methanol) cannot be measured by our current instrumentation, we expect its ^{13}C –D clumping signal is not much different from that of CH_4 . This is because most industrial acetate (the likely source of sodium acetate used in this study) is produced from the high temperature (150 to 200°C) catalytic reaction of methanol and carbon monoxide (*e.g.*, Eby and Singleton, 1983). Industrial methanol, in turn, is produced from carbon monoxide, carbon dioxide, and hydrogen at high temperature (typically 200 to 300°C) (Cheng, 1994). Based on theoretical calculations from modeled vibrational frequencies, Wang et al. (2015) reported that the equilibrium ^{13}C –D clumping of simple carbon compounds have a relatively narrow range of clumped isotope effects from +5.9 to +6.2‰ at 25°C for the molecules studied (methane, methanol, formaldehyde, formic acid, methanethiol, acetic acid). This indicates that the ^{13}C –D clumped isotope effect is not sensitive to detailed bonding environments (also see Piasecki et al., 2016, for *n*-alkane). This is reasonable considering that the ^{13}C –D clumped isotope effect largely originates from a zero point energy shift (ΔZPE) associated with the C–H stretching vibration frequency at around 3000 cm^{-1} ; bending vibration is much lower in energy (1350 cm^{-1}) and the ΔZPE is relatively small (Whitehill et al., 2017). Therefore, we estimate that the acetate or methanol may thus carry a $\Delta^{13}\text{CH}_3\text{D}$ signal of 1.6 to 3.1‰ assuming the near-equilibrium reaction between 150 and 300°C. If we assume that acetate was produced at the lower end of this temperature range (150°C) and methanol was produced at the higher end (300°C), this would correspond to equilibrium values of 3.1‰ for acetate vs. 1.6‰ for methanol. It is possible that this discrepancy could explain part, but not all of the difference between the $\Delta^{13}\text{CH}_3\text{D}$ values of the acetoclastic and methylotrophic cultures. Nonetheless, experiments with various sources of acetic acid,

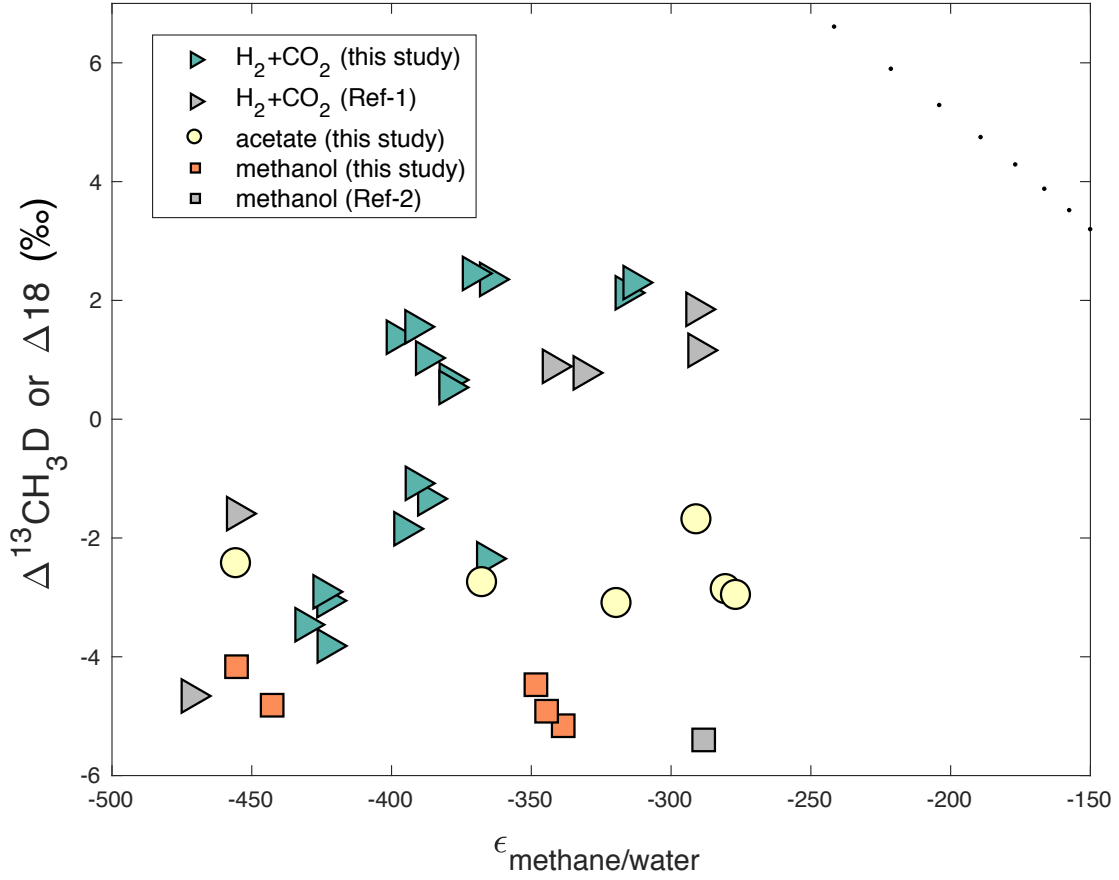


Figure 2-9: The values of $\Delta^{13}\text{CH}_3\text{D}$ (or Δ_{18}) of methane plotted against D/H fractionation factor between methane and water ($\epsilon_{\text{methane/water}}$) for methanogen cultures grown on different substrate. All culture data previously published with corresponding $\delta\text{D}_{\text{H}_2\text{O}}$ values is included for reference (grey symbols). Ref-1 refers to Stolper et al. (2015) and Ref-2 refers to Douglas et al. (2016). Black dots represent isotopic equilibrium (Wang et al., 2015).

natural and/or synthetic, will be needed to constrain the degree of non-equilibrium signals in ^{13}C -D bonds of acetic acids.

Another possibility is that the observed anti-clumped signal originates during the addition of the last hydrogen atom of methane. By applying Equation 2.28, the $\Delta^{13}\text{CH}_3\text{D}$ value of methane (ΔCH_4) is expected to carry $\frac{3}{4}$ signal from methyl precursor ($\text{CH}_3\text{-CoM}$, δCH_3), and $\frac{1}{4}$ from the last H added:

$$\Delta\text{CH}_4 = \frac{3}{4}(\ln \gamma_s + \Delta\text{CH}_3) + \frac{1}{4}\ln \gamma_p \quad (2.29)$$

where γ_s and γ_p are kinetic clumped isotope effects for the secondary and primary D-addition (Equations 2.13 and 2.14). Here, the non-linearity bias in Equation 2.28 does not apply, because ^{13}C isotope effects for secondary and primary processes will be nearly identical (*i.e.*, $\delta^{13}\text{C}_\text{A} - \delta^{13}\text{C}_\text{B} \simeq 0$).

What are the values of γ_S and γ_P ? Applying transition state theory, Whitehill et al. (2017) presented a detailed analysis of the kinetic clumped effect for the gas phase oxidation of methane by the OH radical. We use their framework to make an approximate inference for this study. According to Whitehill et al. (2017), clumped isotope effects can be explained by the difference of zero-point energy shifts (ΔZPE) between $^{13}\text{CH}/^{12}\text{CH}$ and $^{13}\text{CD}/^{12}\text{CD}$. For methane, the ΔZPE for $^{13}\text{CH}_4$ vs. $^{12}\text{CH}_4$ is 29.8 cm^{-1} and that for $^{13}\text{CH}_3\text{D}$ vs. $^{12}\text{CH}_3\text{D}$ is 31.9 cm^{-1} . The difference between the two, $\Delta\Delta ZPE$ of 2.1 cm^{-1} , is the origin of clumped effect (Whitehill et al., 2017). Transition state for the last step of methanogenesis involves a methyl radical intermediate with trigonal planar geometry (Scheller et al., 2013; Wongnate et al., 2016). We estimated the $\Delta\Delta ZPE$ of the methyl radical of 2.1 cm^{-1} using unrestricted MP2 with basis set aug-cc-pVQZ basis set. $\Delta\Delta ZPE$ for reactant ($\text{CH}_3\text{-S-CoM}$, approximated by methylthiol, Wang et al. (2015) and transition state (approximated by the methyl radical) suggests that there is little kinetic clumped effect for the secondary reaction such that γ_S is expected to be close to unity. Whitehill et al. (2017) also showed that the imaginary frequency and tunneling terms do not produce clumped effects (within transition state theory), although the Wigner tunneling correction used in the study is highly approximated.

For the primary reaction, the C-D bond becomes loose at a transition state, contributing a smaller $\Delta\Delta ZPE$ between 0 and 2.1 cm^{-1} (it is 0.5 cm^{-1} for $\text{CH}_3\text{-D-OH}$ transition state, Whitehill et al. (2017). For the bond forming reaction, $^{13}\text{C-D}$ is slightly preferred but not as much as equilibrium ($\gamma_P < 1.006$). However, the anti-clumped effect ($\gamma_P < 1$) is unlikely because it requires the $\Delta\Delta ZPE$ to have the opposite sign (*i.e.*, smaller ΔZPE for $^{13}\text{CD}/^{12}\text{CD}$ than $^{13}\text{CH}/^{12}\text{CH}$) at the transition state. Thus, based on this transition state model, γ_P could take any value between 1.000 and 1.006.

For example, if $\ln(\gamma_S)$ and ΔCH_3 are 0.0 and 1.6 ‰, respectively, the value of $\ln(\gamma_P)$ of -13 ‰ is required to explain the observed $\Delta^{13}\text{CH}_3\text{D}$ value in the acetate culture of $\sim -3\text{‰}$ (Equation 2.29). The magnitude of this kinetic anti-clumped effect has been implicated by Wang et al. (2015) as well as Stolper et al. (2015) to explain the observed values for natural samples. Low $\Delta^{13}\text{CH}_3\text{D}$ values for methanol cultures (-4.2 to -4.9 ‰) suggest that the methyl group of methanol may carry lower $\Delta^{13}\text{CH}_3\text{D}$ values.

Another possibility is that the anti-clumping effect for methanol cultures is related to the faster growth rate of methylotrophic compared to acetoclastic and hydrogenotrophic cultures. Cultures were incubated until they had produced enough methane for our analyses ($>5\text{ mL}$, STP) which was reached over the course of different incubation times as noted in Table 2. Although exact experiment durations and methane concentrations were not measured for all experimental setups, we report estimated growth times and volume of methane produced for Set 2 cultures (Table 2-2). Despite some measurement uncertainty, it appears methanol cultures generated methane more quickly. The effect of varying growth rate is speculative for these experiments, but quantifying the effect of growth rate is something

that should be investigated further in future work.

A series of studies suggest the potential importance of quantum mechanical tunneling in some H transfer reactions, in particular, during the oxidation of alcohol by dehydrogenase (Cha et al., 1989; Klinman and Kohen, 2013). Experimental evidence that supports tunneling includes the observed mass-independent fractionation among H/D/T, departure from the rule of geometric mean for multiply-deuterated (clumped) substrates, rate enhancement at low temperatures (Klinman and Kohen, 2013; Srinivasan and Fisher, 1985; Amin et al., 1988; Huskey, 2006), and the anti-clumped CH₂D₂ abundance (Young et al., 2017). The large anti-clumped effect may be due to the tunneling effect, which is only approximated in the above transition state model. The test for this hypothesis would include *in vitro* enzyme assay experiments similar to Scheller et al. (2013), or high-level quantum mechanical modeling with accurate geometry and potential energy surface at the key transition state (*e.g.*, Chen et al., 2012; Klinman and Kohen, 2013; Wongnate et al., 2016).

2.5.5 Non-equilibrium vs equilibrium $\Delta^{13}\text{CH}_3\text{D}$ signals of methane in the environment

This study corroborates previous studies (Stolper et al., 2015; Wang et al., 2015; Douglas et al., 2016; Wang et al., 2016; Young et al., 2016, 2017) and demonstrates that non-equilibrium (*i.e.*, kinetic) clumped methane isotopologue signals are common for methane produced by microbial methanogenesis in laboratory cultures (Figure 2-1, Figure 2-9).

Our results are consistent with kinetic signals in methane sampled from freshwater environments (*e.g.*, swamps and lakes), where acetoclastic methanogenesis, as opposed to hydrogenotrophic methanogenesis, is thought to be the dominant source of methane (*e.g.*, Conrad, 2005; Ferry, 2010). Nonetheless, our results demonstrate that low $\Delta^{13}\text{CH}_3\text{D}$ values are consistently out of equilibrium in batch culture and in environmental samples, but that this is not necessarily dependent on the metabolic pathway (hydrogenotrophic, methylo-trophic, or acetoclastic). Thus, $^{13}\text{CH}_3\text{D}$ isotopologue compositions cannot be used alone to resolve which methanogenic pathway is dominant in the environment.

Although dominantly microbial in origin, methane in marine environments (*e.g.*, pore water and hydrate) tends to carry equilibrium or near-equilibrium $^{13}\text{CH}_3\text{D}$ abundances (Stolper et al., 2015; Wang et al., 2015). This is also corroborated by δDCH_4 values that are relatively constant at $-180 \pm 10\text{‰}$ (*e.g.*, Whiticar, 1999; Okumura et al., 2016), which is close to the expected value for methane in equilibrium with seawater (with $\delta\text{D} \simeq 0\text{‰}$) (Horibe and Craig, 1995). It was suggested that "slow" methanogenesis under small thermodynamic drive (low environmental H₂ concentration) would produce near-equilibrium clumped isotopologue signals (Wang et al., 2015; Stolper et al., 2015). Oxidative cycles of methane also modify clumped isotopologue signals of environmental methane. The aerobic oxidation of methane by *Methylococcus*, for example, is characterized by kinetic clumped

isotope effects ($\gamma \simeq 1$), and residual methane can exhibit strong non-equilibrium signals (Wang et al., 2016). In contrast, anaerobic oxidation of methane (AOM), thought to be the reverse process of methanogenesis (Hallam et al., 2004; Scheller et al., 2013; Wang et al., 2014), has been shown to operate with significant back-flux (Holler et al., 2012) and produce near-equilibrium carbon isotope fractionation (*e.g.*, Yoshinaga et al., 2014). Methane cycling involving AOM and methanogenesis likely promotes near-equilibrium isotopologue distributions, in particular, in marine environments, and may contribute to near-equilibrium δD as well as $\Delta^{13}CH_3D$ signals. Conversely, bulk δD and $\Delta^{13}CH_3D$ values of methane may primarily reflect kinetic versus equilibrium signals and the rate of methanogenesis more so than metabolic pathways.

2.6 Conclusions

This study reports the $\Delta^{13}CH_3D$ systematics of microbial methane produced by pure cultures of methanogens. Our results show that the $\Delta^{13}CH_3D$ signals are not directly pathway dependent, as cultures of *M. barkeri* and *M. mazei* grown on acetate, methanol, and $H_2 + CO_2$ all yield methane that is depleted in $^{13}CH_3D$, which seems to be characteristic of microbial methanogenesis in near-surface environments (lakes, swamps and ruminants). For mesophilic methanogens, the lowest $\Delta^{13}CH_3D$ values were produced for methane from cultures grown on methanol. Methanol cultures grew faster (incubation time of 3 days) than those grown on acetate or $H_2 + CO_2$ (incubation time of 30 days). Thus, the $\Delta^{13}CH_3D$ values of methane may be related to the rate at which methane is produced rather than to the substrate used. Mesophilic methanogens (*M. barkeri*) produced anti-clumped $\Delta^{13}CH_3D$ values (<0), while thermophilic and hyperthermophilic (*Methanothermococcus* and *Methanocaldococcus*) methanogens produced less kinetic signals.

Experiments with deuterated water or acetate aid in determining the source of hydrogen atoms in methane. The deuterated water experiments confirm that the four hydrogen atoms that form methane in hydrogenotrophic methanogenesis are derived from water. For the acetoclastic culture, 1.6 to 1.9 H atoms are derived from water, whereas only one is required by stoichiometry, suggesting some reversibility and isotope exchange at the methyl precursor. The deuterium spiked experiments also demonstrate that the observed non-equilibrium signals cannot be explained by the mixing of two pools of C-H bonds (*e.g.*, from methyl group of acetate and one C-H bond formed during acetoclastic methanogenesis). The production of low $\Delta^{13}CH_3D$ values independent of the methanogenic pathway, suggests, although not exclusively, that the most of the kinetic signal is produced during the enzymatic reactions common in the three methanogenic pathways, such as the reduction of methyl-coenzyme M.

CHAPTER 3

EVOLUTION OF METHYLOTROPHIC METHANOGENESIS AND POSSIBLE IMPLICATIONS FOR BIOGEOCHEMICAL CYCLES IN PROTEROZOIC AND PHANEROZOIC OCEANS

3.1 Abstract

Recent work employing molecular clock dating techniques has shown that one major group of methanogens acquired the ability to use acetate as a methanogenic substrate via horizontal gene transfer (HGT), which likely precipitated major changes to the carbon cycle during the end of the Permian. It is hypothesized that another major pathway of methanogenesis, methylotrophic methanogenesis also evolved via HGT, relatively recently. We explore the emergence of methylotrophic methanogenesis by studying the phylogenetic history of substrate-specific methyltransferase enzymes and corrinoid proteins. The relationship between methyl-corrinoid sequences illustrates substrate specificity with patterns of gene duplication and loss. While the exact evolutionary history of these substrate-specific methylotrophic genes is unknown, the phylogenetic patterns presented here are consistent with the hypothesis involving HGT from bacteria to archaea. Moreover, the relationship between taxa that possess methyl-corrinoid genes suggests strong ties to the marine environment. Methylotrophic methanogens may be unique in their substrate links to the evolution of marine eukaryotic algae, purple sulfur bacteria, carbon and sulfur cycles, and the chemical inventory of the ocean. This work forms a more dynamic view of the contributions of microbial methane production by the major pathways of methanogenesis over the last four billion years and may yield further insight into the role of methylotrophic metabolisms in marine systems.

3.2 Introduction

3.2.1 Microbial Methanogenesis

The origin of methane reflects a dynamic biogeochemical history (Hinrichs, 2002). While the earliest sources of methane production on Earth were abiotic (thermogenic and abiogenic), over time, microbial methane has come to dominate the global inventory (Whiticar, 1990). This shift coincided with major events in the evolution of both life and Earth (Figure 3-1) (Rothman et al., 2014). The greatest source of methane today is the microbes that live in anoxic environments such as swamps, sediments, rice paddies, and ruminant tracts via methanogenesis (Klapp et al., 2010). As one of the earliest metabolic pathways to evolve (Baptiste et al., 2005), microbial methanogenesis has played an important role in the biogeochemical cycling of carbon, not only in the modern environment, but also since the evolution of methanogenic metabolism (Rothman et al., 2014) over 3.46 billion years ago (Ga) (Ueno et al., 2006; Wolfe and Fournier, 2018). Although often thought of as an ancient metabolism, methanogenesis nonetheless utilizes a diverse set of relatively young sub-pathways and substrates, which have continually evolved in response to changing planetary conditions (Rothman et al., 2014).

The emergence, evolution, and proliferation of microbes have each fundamentally reshaped, and been shaped by, Earth's chemical inventory. Recent work dating the evolution of one specific pathway, acetolactic methanogenesis, suggests that a horizontal gene transfer (HGT) event expanded the metabolic capability of methanogens to utilize acetate as a substrate; this was associated with a pronounced increase in methane production during the end Permian (Rothman et al., 2014). Developing a better quantitative and qualitative understanding of the role of microbes in methane cycling is critical for defining the distribution of carbon on Earth over the last 4 billion years, and for predicting how this system may be affected by future changes. However, not all methanogenic pathways are well understood in the context of their evolutionary history or their environmental links. The goal of this chapter is to use molecular evidence from genetic sequence databases to inform a better understanding of the possible biogeochemical links of, and evolutionary events involving, methylotrophic methanogens during the Proterozoic and Phanerozoic eons. This is expected to compliment other lines of evidence, such as the geochemical record, and inform future work to understand the biological impact on the evolution of the carbon cycle.

Microbial methanogenesis is an anaerobic metabolism restricted to Euryarchaeota (Baptiste et al., 2005; Laso-Pérez et al., 2016; Spang et al., 2017). Methanogens produce methane as a catabolic end product of three substrate-specific pathways that feed methyl groups to a conserved terminal pathway (Figure 3-2) (Thauer, 1998). Methylotrophic methanogenesis, the metabolic pathway employed by Methanosarcinales and Methanomassiliicoccales, uses methyl compounds (methanol, methyl amines, dimethylsulfide (DMS), or methylthiols) as substrates (Baptiste et al., 2005; Penger et al., 2012) (Figures 3-3 and 3-4). Methanobacteriales (including the two gut microbes *Methanobrevibacter smithii* and *Methanosphaera stadtmanae*) and *Methanomassiliicoccus* can also use methanol in the presence of hydrogen (Dridi et al., 2012; Fricke et al., 2006).

Multiple lines of evidence indicate that microbial methane was formed early in the Archaean (Battistuzzi et al., 2004; Rothman et al., 2014; Ueno et al., 2006; Wolfe and Fournier, 2018). Geochemical evidence suggests that putatively microbial methane was produced at least 3.46 Ga (Ueno et al., 2006), and molecular clock evidence suggests that methanogens diverged at least 3.51 Ga (Wolfe and Fournier, 2018). It is striking that over the last 3.5 billion years of Earth history, methanogenesis has remained one of the most highly conserved metabolic strategies, and that this pathway has never been horizontally transferred to another group of microbes (Fournier et al., 2009). Nonetheless, methanogens have continued to evolve in ways that diversify their metabolic capabilities and impact their role in major biogeochemical cycles. Until only very recently, it was thought that all methanogens belonged to the phylum Euryarchaeota. Compelling but still incomplete genomic evidence suggests that an uncultured Archaeon, Bathyarchaeota, (Evans et al., 2015) possess homologs of the genes necessary for methanogenesis, and that this group may be placed outside

of known archaeal clades. Thus, the evolution of microbial methanogenesis may even pre-date Euryarchaeota, and be an ancestral archaeal metabolism. This also suggests that a vast diversity of microbes may remain uncultured, particularly from remote environments; thus, the molecular record may hold vital clues to illuminating the evolution of this metabolism (Sogin et al., 2006).

3.2.2 Horizontal Gene Transfer and Molecular Clock Dating

Horizontal gene transfer (HGT) has played an important role in the evolutionary history of many metabolic pathways, including methanogenesis (Fournier et al., 2009). While vertical inheritance passes genetic information from parent to offspring, HGT passes information between lineages, even if they are very distantly related. The importance of HGT in microbial evolution has become increasingly appreciated (Gogarten, 1995; Gogarten and Townsend, 2005; McInerney et al., 2008). For example, Haloarchaea acquired genes from a combination of bacterial donors which facilitated a novel pathway for carbon assimilation (the methylaspartate cycle) and access to a new metabolic niche (Khomyakova et al., 2011; Soucy et al., 2015). HGT has also been shown to be a significant mechanism driving antibiotic resistance in bacteria (Gyles and Boerlin, 2014). HGT events are detected by comparing phylogenies inferred from highly conserved protein sequences (and thus more likely reflecting the species tree) to those of individual genes. Topological conflicts between gene and species trees are candidate HGT events (Soucy et al., 2015). A variety of computational tools exist to detect and evaluate these events (Dos Reis et al., 2015; Fournier et al., 2015; Ravenhall et al., 2015; Zhu et al., 2014).

Molecular clock techniques can be used to date the evolution of plants and animals by calibrating phylogenetic trees with fossil evidence (Peterson et al., 2004). The underlying principle of such analyses is that not all genes within an organism evolve with the same rates or under the same evolutionary mechanisms. A phylogenetic tree is used to provide relative divergence distances, and a molecular clock estimates the rate of evolution in the tree. Absolute timing of the molecular clock can be calibrated using independently known dates of characteristic fossils, similar to how index fossils are used to date stratigraphic columns in geologic settings (Donoghue and Benton, 2007; Hedges and Kumar, 2003; O'Reilly et al., 2015). Therefore, hypotheses regarding evolutionary relationships and the timing of evolutionary events can be tested using the most parsimonious explanations. However, most microbial life does not leave behind fossils with which to date divergences. Consequently, new molecular clock techniques have been developed to use extant gene and protein sequences to estimate divergence times of microbial lineages (Donoghue and Benton, 2007; Dos Reis et al., 2015; Drummond et al., 2006; Knoll, 2017; Parfrey et al., 2011; Schirmer et al., 2015). HGT, coupled with new molecular tools, has been proposed as one way to integrate the timing of events in microbial evolution with stronger confidence and higher precision

(Dos Reis, 2018; Knoll, 2017; Magnabosco et al., 2018; Wolfe and Fournier, 2018).

The methanogenic pathway presents a useful metabolism for detecting horizontal gene transfer events, because this pathway is one of the oldest, most highly conserved, and unified metabolic systems. There is no known HGT of a methanogenic pathway to a non-methanogenic lineage (Gribaldo and Brochier-Armanet, 2006) except in the case of the recently described *Verstraetearchaeota* (Spang et al., 2017). This may be because methanogenesis requires many phylogenetically distinct genes and complex cofactors and carrier proteins which safeguard against HGT. Moreover, all methanogens appear to have inherited one core pathway from a distant common ancestor (Fournier et al., 2009). This provides a characteristic signal of methanogenic metabolism in the genetic record and easy detection for HGT to and between methanogens. It is hypothesized that both acetoclastic methanogens and methylotrophic methanogens bear signatures of HGT, but with very different scenarios.

The methylotrophic pathway involves the transfer of a methyl group from the substrate (*e.g.*, methanol, methylamines, methylthiols, or dimethylsulfide) via two methyltransferase enzymes and a corrinoid protein to coenzyme M, and then follows the pathway common to all methanogens (Figures 3-3 and 3-4) (Ferry, 2010; Sauer and Thauer, 1998). Each substrate for this pathway has a unique, corresponding corrinoid protein and methyltransferase enzyme (Fournier et al., 2009). The only other organisms with the methyltransferase gene are Bacteria which are not capable of methanogenesis. Interestingly, this pathway requires the noncanonical amino acid pyrrolysine (Pyl), which is used by no other proteins (Fournier et al., 2009). Pyl is rare in extant life, and is utilized only by a subset of methanogens for the synthesis of methanogenesis enzymes from methylamines. Methanogens likely acquired this protein via HGT (Fournier et al., 2009) of a novel aminoacyl-tRNA synthetase (Fournier et al., 2009). The donor lineage was likely ancient but remains unknown. It is further hypothesized that there may be an evolutionary relationship between Pyl and the proteins used in the methyl-corrinoid pathway of methanogenesis (Borrel et al., 2014; Ferguson et al., 2009), but this is an avenue which requires further investigation.

Methanosarcinales, Methanomassiliicoccales, and *Verstraetearchaeota* are the only orders that use the methyl-corrinoid pathway (Figure 3-5) Vanwonterghem et al. (2016); Spang et al. (2017). Because these orders of methanogens are not closely related but share this very specific cellular machinery, it is hypothesized that the corrinoid proteins and methyltransferases were distributed between these lineages via HGT (Deppenmeier et al., 2002). Moreover, these proteins are present in multiple copies and in many cases are related to one another. Thus, we hypothesize that there was a complex evolutionary history of gene duplication that must be unraveled to understand how and when these pathways evolved. An extensive assessment of the phylogeny of methylamine methyltransferase protein families has never been performed, but would allow for the testing of evolutionary hypotheses, and may present strategies for future dating of these events.

3.2.3 Microbial Links in Earth’s Biogeochemical Cycles

It has been hypothesized that the chemical inventory of the early Earth influenced the requirements for microbial methanogenesis. This hypothesis is of particular interest in explaining metal specificity, because methanogenesis has been cited as one of the most metal-rich enzymatic pathways in microbial biochemistry (Zerkle et al., 2005). It is well established that nickel is required for methyl-coenzyme M (CoM) reductase, which is utilized by all pathways of methanogenesis (Scheller et al., 2010; Thauer, 1998). But methanogenic pathways also require many other metals as cofactors, including Fe, Ni, Co, Mo, W and Zn. Studies have shown that when Fe, Ni, and Co are limited, methanogenesis is reduced (Glass and Orphan, 2012; Lessner, 2009; Sauer and Thauer, 1998; Scheller et al., 2010). Anaerobic marine ecosystems likely selected for the use of elements that were more readily available (David and Alm, 2011; Dupont et al., 2010, 2006; Glass and Orphan, 2012). It has been suggested that the Archaean ocean was Cu- and Zn-poor, and thus Cu- and Zn-dependent metabolisms evolved later (Dupont et al., 2010). This hypothesis would explain the lack of Cu-dependent anoxic methanogenesis enzymes (Glass and Orphan, 2012). As Cu became more bioavailable in an oxygenated world, aerobic methanotrophs may have been forced to depend on previously scarce and unused metals (David and Alm, 2011; Dupont et al., 2010, 2006; Glass and Orphan, 2012). Zn is hypothesized to have become available later in Earth’s history as well, which corresponds to the characteristic presence of Zn-binding proteins in later-evolving Eukaryotic metabolisms (Dupont et al., 2010). Modern genomes may be imprinted by, and retain signals of, such biogeochemical events (David and Alm, 2011). For example, it has been shown that the bioinorganic chemistry of ancient oceans has shaped the evolution of cyanobacterial requirements (Saito et al., 2003). Thus, it stands to reason that the presence of necessary metals or key substrates in these marine environments may have also enabled the diversification of Archaea and influenced the resulting metabolic pathways.

Microbes have played an important role in the evolution of the Earth’s carbon and sulfur cycles, and the genetic record of extant organisms can expose traces of such biogeochemical inputs (David and Alm, 2011). The evolution and proliferation of methanogenic metabolisms has obvious consequences for the carbon cycle, as methane is one of its key intermediates. For instance, HGT was implicated in the assembly of the methanogenic pathway from Bacteria to acetoclastic *Methanosarcina*; this resulted in an increase in methane production during the Permian (Fournier et al., 2009; Rothman et al., 2014). Additionally, methylotrophic methanogenesis links the carbon and sulfur cycles to the three domains of life in the marine ecosystem through the metabolism of the substrate dimethylsulfide (DMS) (Curson et al., 2011) (Figure 3-6).

Although flux estimates are debated, some estimates indicate that DMS is responsible for 50-60% of the reduced sulfur flux to the atmosphere, 95% of which is derived from

the oceans (Stefels et al., 2007). The major source of DMS in the ocean is the bacterial degradation of dimethylsulfoniopropionate (DMSP) to DMS (and to a much lesser extent, algal degradation) (Alcolombri et al., 2015; Li et al., 2014; Todd et al., 2007; Yoch, 2002). Additionally, a small portion of DMS is produced by the bacterial metabolism of methanethiol, and bacteria can also reduce dimethyl sulfoxide (DMSO) to DMS. Bacteria can utilize both DMS and DMSP as a carbon or energy source (Green and Hatton, 2014). Although its exact purpose is unclear, DMSP is produced by photosynthetic Green Algae (*e.g.*, Chlorophytes), Dinophyceae (Dinoflagellates), Prymnesiophyceae (Prymnesiophytes, including Coccolithophores), Chrysophyceae (Golden Algae) Bacillariophyceae (diatoms), and a few Angiosperms that live by the sea, likely for osmoregulatory purposes (Curson et al., 2011; Kettle et al., 1999; Kettles et al., 2014; Panos and Jones, 2013). In full, this production results in estimated ocean DMSP concentrations of 50-400 mM, which can account for a significant portion of the total organic sulfur cellular material (Stefels et al., 2007) and thus supplies much of carbon and sulfur that Bacteria require (Vila-Costa et al., 2014). However, the algal enzyme responsible for DMSP degradation to DMS (DMS lyase) was only recently identified (Alcolombri et al., 2015), and the genes responsible for DMSP synthesis in algae remain poorly understood (Kettles et al., 2014).

Almost all Bacteria capable of catabolizing DMSP to DMS have been identified from the phylum Proteobacteria. These taxa include mostly marine Gammaproteobacteria and the abundant Alphaproteobacteria (*Roseobacters*), and to a lesser extent Beta-, Delta-, and Epsilonproteobacteria (Curson et al., 2011). Purple Nonsulfur Bacteria (*e.g.*, *Rhodospirillum rubrum* and *Rhodocyclus tenuis*) have also been experimentally shown to produce DMS during photosynthesis (Agalidis et al., 1997; Munk et al., 2011). The major genes associated with DMS and DMSP metabolism, as well as their source organisms and evolution dates, are summarized in Tables 3-2 and 3-3, respectively.

Phytoplanktonic degradation or use of DMSP and production of DMS provides an important link in the sulfur and carbon cycles, impacting many levels of the microbial food web (Vila-Costa et al., 2006). DMS is often implicated in the well-known, albeit controversial, "CLAW" hypothesis. This hypothesis posits that DMS from oceanic phytoplankton leads to the formation of aerosol particles that act as cloud condensation nuclei in the marine boundary layer, driving a subsequent increase in cloud formation, albedo, and temperature change (Charlson et al., 1987). Contention over this hypothesis stems from a lack of observable evidence and indication that the system is more complex than previously described (Ayers and Cainey, 2007; Quinn and Bates, 2011). Nonetheless, it has been suggested that the microbial role in this process may actually be much greater than that described by the CLAW authors (Green and Hatton, 2014).

While DMSP production and degradation is a topic of active research, less attention has been paid to DMS consumers. A large portion of the DMS produced in the ocean (as

much as 50%) does not make it into the atmosphere (Vila-Costa et al., 2006). This is partly due to metabolic utilization. Although DMS metabolism and its ecological distribution are not entirely constrained, methylotrophic methanogens have been shown to grow on DMS in both marine and freshwater systems (Lomans et al., 2002). Methylotrophic methanogenesis may play a more dominant role in the SRZ of marine sediments. It has been shown that decreasing total organic carbon and the depth of the sulfate reduction zone in marine sediments may lead to relatively more methylotrophic microbial activity. It is hypothesized that this increased activity is because DMS is a noncompetitive substrate in this environment (Summons et al., 1998; Zhuang et al., 2016). Figure 3-6 illustrates a simplified schematic of a possible pathway for DMS production and use. Evolutionary ages of DMS-consuming methanogens may correspond with the emergence of DMSP production by Algae and DMSP to DMS breakdown by Bacterioplankton. Dating other corrinoid pathway proteins may also be relevant to this timing. It is hypothesized that DMSP became abundant in the environment relatively recently, so the capacity to produce and catabolize this molecule must also have evolved relatively recently (Curson et al., 2011). Because algal DMSP producers appeared in the Proterozoic, it is possible that methanogens acquired the ability to utilize DMS as a substrate later than this if they diversified in the marine system, although precise dates are unknown. Preliminary data indicate that extant Purple Nonsulfur Bacteria diverged around 600-1000 Ma, and that they diverged from other Alphaproteobacteria around 850-1200 million years ago (Fournier, Pers. Com.). Even at these low levels of precision, however, these results are consistent with the hypothesis that DMS cycling evolved in the Neoproterozoic.

3.2.4 Summary and Objectives

The goal of this chapter is to investigate the phylogenetic history of genes associated with the methyl-corrinoid pathway of methanogenesis. Can a genomic approach can reveal the evolution substrate-specificity and links to the environment? It is hypothesized that evolution of the methyl-corrinoid pathway may be connected to the marine system and tied to ocean chemistry due to metabolic dependence on substrate availability and even metal specificity. Therefore, we seek to investigate the phylogenetic history of methyl-corrinoid proteins to assess whether the capability to use specific methyl compounds was acquired via HGT; in what group of methanogens this pathway originally evolved; whether the evolution of this pathway is linked to chemical inventory constraints (*e.g.*, methylthiols, methanol, methylamines, DMS or trace metals); or whether corrinoid-pathway methanogens coevolved with other taxa implicated in DMSP-DMS metabolism (*e.g.*, Green Algae or Purple Nonsulfur Bacteria).

3.3 Methods

3.3.1 Identification of Genes of Interest and Test for Homology

Methyltransferase and corrinoid proteins were identified by conducting a literature review of the methylotrophic pathway, and by querying databases including NCBI and MetaCyc. These proteins are listed in Table 3-4. The identities of sequences of interest were further assessed using NCBI BLASTp. Additionally, to test whether this set of putatively substrate-specific proteins was homologous, proteins were aligned in pairs with NCBI's BLAST2. Sequences that produced significant hits, defined as having an E-value $<10^{-5}$, were considered homologous. Homologous protein families were then aligned to each other for subsequent analyses.

3.3.2 Taxon Sampling

We queried The National Center for Biotechnology Information (NCBI) nonredundant (nr) database using the protein Basic Local Alignment Search Tool (BLASTp) for homologs related to the mtaC methyltransferase corrinoid protein (Q46EH4). We assessed the similarity of homologs (E-value $<10^{-5}$) and downloaded the complete sequences of all top archaeal hits from NCBI (~ 700 taxa). Substrate specificity was determined based on annotation and sequence identification. We further used BLASTp to identify significantly similar bacterial sequences for all major substrate groups on the tree.

3.3.3 Sequences and Alignments

Sequences were aligned using MUSCLE (Edgar, 2004).

3.3.4 Phylogenetic Analyses and Model Selection

Single gene trees were inferred using RaxML v1.8.9 using the PROTGAMMALGF substitution model (Stamatakis, 2006) and 100 bootstrap replicates. The best-fit model for the amino-acid substitution was assessed using ProtTest (Darriba et al., 2017).

3.4 Results and Discussion

3.4.1 Single Gene Trees and Homologous Proteins

Single gene trees were built for methyltransferase and corrinoid proteins. Figures D-1-D-12 illustrate these gene trees with midpoint rooting. The red is indicative of putatively methanogenic methyl-corrinoid sequences. Methyltransferases and corrinoid proteins were

compared to each other using BLAST2, and produced significant hits for all sequence combinations, supporting the hypothesis that they are homologous. One corrinoid protein, mtaC, was selected for construction of a more extensive gene tree, serving as a representative homolog of the other substrate-specific corrinoid proteins.

3.4.2 Homologous Corrinoid Proteins

The mtaC gene tree (Figure 3-7, Taxa on Table 3-4) depicts the top archaeal BLAST hits as well as the top bacterial hits for all major substrate groups. Major groups of substrate-specific sequences clearly group together, falling into the following categories based on annotation: methionine, dimethylsulfide, trimethylamine, dimethylamine, methylamine, methylthiol, methanol, or unidentified substrate.

3.4.3 Rooting the Methyl Corrinoid Tree

Bacteria were included 1) to identify HGT to and from bacteria involving the evolution of the corrinoid proteins in methanogens and 2) to help root the tree, identifying divergent outgroups to all the methanogenic corrinoid proteins. The mtaC tree contains every significant taxon hit within archaea. Taxa were put into groups identified by both their function and clustering in an unrooted tree. Visual inspection of the tree suggests that the most distantly-related group is likely the methionine cluster. Bacterial and archaeal sequences with a representative of each group were used to perform a BLAST search; each bacterial hit that was more closely related to the query sequence than to any other archaeal sequence already in the tree was added to the dataset. All selected taxa were then realigned to form the final mtaC tree (Figure 3-7). Adding bacteria enabled the identification of (1) possible bacterial clades that might be nested within an extant archaeal corrinoid protein group (due to HGT), and (2) likely bacterial outgroups of any more deeply-diverging homologs. Additionally, adding bacteria seems to support the rooting and placement of the methionine cluster as one of the more deeply-branching clades, supporting the observation based on the grouping and distribution of bacterial taxa in the unrooted tree.

3.4.4 Corrinoid Protein Tree Topology and Substrate Specificity

The protein mtaC is a corrinoid protein harboring subunit of the methyltransferase protein, typically referred to as a methanol-specific methyl corrinoid protein (Dong et al., 2017; Kröninger et al., 2017; Sauer and Thauer, 1998; Tallant and Krzycki, 1997). Here we refer to it simply as mtaC. Annotation and sequence assessment with NCBI reveal regions of distinct substrate specificity on the mtaC single gene tree (Figure 3-7). Some regions of the tree lack substrate designations because their review was beyond the scope of this work, due to incorrect annotations in the source database. While there are some informative regions of the

tree, there are may be regions of this tree with missing, misannotated, or still undetermined substrate specificity, which will be discussed later in this chapter. Each substrate-specific region of the tree appears to contain some areas consistent with published species-level relationships, suggesting that this tree preserves some patterns of vertical inheritance with many areas of suspected HGT. This should be investigated in future work.

Methionine, Dimethylsulfide, and Methylthiol

Methylated thiol substrates such as dimethylsulfide (DMS) and methylmercaptopropionate (MMPA) have been identified as being precursors to microbial methane (Tallant et al., 2001). Methionine itself is not considered to be a substrate for archaeal methanogenesis. However, these methylated thiols can be produced from the degradation of compounds such as methionine. Thus, there may be a relationship between these more recently identified substrates and the methionine-annotated protein, which may also explain this region of the tree. The DMS-specific pathway of methanogenesis requires two proteins: *mtsA* and *mtsB*. It has been shown that these proteins are homologous to the cobalmin binding domains of methionine synthase (Tallant et al., 2001). It may be that these genes are most closely related to DMS substrate utilization, or they may be nonspecific. This also makes sense given that at least one of the two methionine-annotated sequences are grouped relatively closely to other DMS-identified sequences. The other group is on a considerably long branch, indicating its placement is possibly biased by long-branch attraction (Philippe et al., 2005).

Regions on the tree that are annotated as methionine-specific, or proposed to be associated with this substrate, contain a few notable members. This group consists of methanogens, including *Methanomassillicococcales*, and *Methanoculleus*. It also includes other nonmethanogenic archaea such as and Bathyarchaeota (a proposed butane oxidizer) *Nitrososphaera* (an ammonia oxidizer), *Nitrosopumilus* (a Thaumarchaeote common in seawater), *Lokiarchaeum*, and *Candidatus* 'Syntrophoarchaeum'. Interestingly, the latter genus has been proposed to contain genes similar to methyl-coenzyme M and butyl-coenzyme M, which suggests that this thermophilic Archaeon can activate butane via alkyl-coenzyme M formation (Laso-Pérez et al., 2016). The finding that *Candidatus* 'Syntrophoarchaeum' is closely related to *Methanomassillicococcales* is consistent with its phylogenetic placement based on the highly conserved McrA gene in the literature (Laso-Pérez et al., 2016). Moreover, like *Lokiarchaea* isolated from deep marine sediments (Spang et al., 2015), *Syntrophoarchaeum* is present in Guaymas Basin marine sediment. This further supports the hypothesis that marine sediments may be rich in methyl-cycling metabolisms. This is also consistent with the notion that oceanic sediments are ripe for metabolic exploration, and some of these methyl-related enzymes are promising for future work detecting and exploring these metabolic pathways or informing future culturing work. These environments may be more significant and metabolically diverse than previously thought.

Finally, if the tree is rooted with taxa from one of the largest bacterial groups on the tree (*e.g.*, Chloroflexi_bacterium_RBG_16_54_18|hypothetical|OGO33135.1), methionine-annotated regions are the most deeply rooted, grouping most closely to the larger bacterial clades that include members such as *Desulfosporosinus*, *Ignavibacteria*, *Deltaproteobacteria*, and Firmicutes (*e.g.*, *Clostridium*). This placement is consistent with the idea that methanogens acquired some of their genes via HGT from bacteria. Firmicutes, in particular, are implicated in the HGT to acetoclastic methanogens (Fournier and Gogarten, 2008).

Although no sequences in the mtaC gene tree were specifically annotated as "dimethylsulfide" (DMS), when the NCBI protein database was cross-referenced for "dimethylsulfide corrinoid", representative sequences were found in the tree, as illustrated in Figure 3-7. The sequences found by this cross-referencing all form a clade, and many have alternative identical sequence names that are annotated as dimethylsulfide in the NCBI database. This clade includes methanogens such as *Methanosarcina*, *Methanlobus*, *Methanococcoides*, *Methanohalobos*, *Methanohalophilus*, and *Methanosalsum*, and is sister to another smaller clade of sequences that encode enzymes that might be DMS-specific as well.

Although manually assessing and reannotating the thousands of genomes in this study is beyond the scope of this work, some of these closely related taxa (*e.g.*, >SDG04186.1) were assessed by BLAST analysis and by reviewing the literature associated with the protein entries. Despite what may be annotated, the closely related proteins were determined to indeed be part of the corrinoid protein. Moreover, many of the highest hits (most closely related sequences with highest percent identity and lowest E-Values) were actually annotated as "dimethylsulfide," "methyltransferase," or "methionine synthase" (*e.g.*, >AKB35381.1). This result highlights the importance of manually reviewing sequences of interest and the possibility of future work with these data. Moreover, it suggests that the closely-related taxa are indeed also DMS-specific corrinoids, building a case for identifiable and discrete substrate specificity and highlights numerous duplications and transfer events. Furthermore, many of the papers that describe methyltransferase pathway function, distribution, and phylogenetic characteristics never mention the dimethylsulfide pathway, thus future work would be particularly interesting in this area.

BLAST results for taxa within the closely related unidentified substrate region had highest sequence similarity to methyl corrinoid proteins. Some were not specific to DMS—for example, >WP_012194551.1 was annotated either with no additional substrate information or as dimethylamine-specific. This annotation is reasonable, however, as the region encoding this protein is mtbC, consistent with dimethylamine. Other closely related taxa include nonmethanogenic Archaea such as *Hadesarchaea*, of the South-African Gold Mine subsurface (Baker et al., 2016). Based on gene composition analyses, *Hadesarchaea* are thought to be involved in CO and H₂ oxidation, possibly coupled to nitrite reduction to ammonia. It is interesting that *Hadesarchaea* are present on this mtaC gene tree. They

lack the often diagnostic *mcrA* gene and are thought not to be capable of methanogenesis (Baker et al., 2016), but appear to be closely related to *Ca.* 'Bathyarchaeota' and contain many methanogen genes (Laso-Pérez et al., 2016). Also of note, the hyperthermophilic iron-oxidizing Archaeon *Ferroglobus* sits nested within many of Firmicutes and other bacterial taxa. *Ferroglobus*, isolated from a subsurface marine hydrothermal system near Italy (Hafenbradl et al., 1996), was the first Archaeon that was shown to anaerobically oxidize aromatic compounds (Tor and Lovley, 2001). Its position within Bacteria may be indicative of HGT, or a gap that remains in our sampling of marine and subsurface sediment environments.

The DMS-specific region of sequences is closely related to bacteria including firmicutes (*Clostridia*, *Sporomusa*, *Pelosinus*, *Syntrophobotulus*, *Anaerovibrio*, *Desulfosporosinus*, *Proteiniborus*), *Treponema*, and Deltaproteobacteria. It is notable that *Sporomusa* and *Pelosinus*, of the Firmicute class Negativicutes, are thought to have acquired membrane characteristics via HGT. This further supports the hypothesis that HGT is rampant in bacterial lineages (Gupta, 2011). Another member in the DMS-specific region of the tree is *Treponema primitia*, a acetogenic spirochaete that both lives in close association with and contains similar genetic information as methanogens (Graber and Breznak, 2004).

Methylthiol is not a specific substrate of methyltrophic methanogenesis. However, the methylthiol methylcoenzyme M methyltransferase has been shown to mediate coenzyme M methylation for cultures (*M. barkeri*) grown on DMS and MMPA substrates in culture. It was also shown that methanogens possessing the methylthiol methyltransferase may grow on 3-methylmercapto-1-propanol. However, it is noted in these studies that methanogens also required acetate in order for the methylthiol methylcoenzyme M methyltransferase to convert DMS or MMPA to methane (Tallant et al., 2001; Tallant and Krzycki, 1997). Another difference between the methylthiol:CoMmethyltransferase and the proteins mediating CoM-methylation with other methyltrophic substrates is that CoM methylation with either methanol, TMA, or MMA requires three polypeptides rather than the two-subunit protein for methylated thiols (Tallant and Krzycki, 1997).

Based on the sequence annotations for the methylthiol region of the *mtaC* gene tree, it appears that nearly all methylthiol sequences belong to methanogenic archaea: *Methanosarcina*, *Methanomethylophilus*, *Methanomassillicoccus*. A BLAST search of the NCBI database for *Methanosarcina barkeri* WP_048107235.1 methylthiol methyltransferase reveals that the most closely related sequences are also all methylthiol-specific methyltransferase proteins, and that the most closely related sequences are annotated as "methanol" similar to the observed pattern in the *mtaC* gene tree (Figure 3-7). The methylthiol region of the gene tree comparatively smaller. Methanogens may be capable of restricted (requiring the use of acetate) or more flexible (growth solely on methyl thiol substrates) metabolisms using this enzyme, but this pathway does not appear to be as taxonomically-diverse. The environmental implications of these observations would be interesting to explore further, by assessing

the presence, distribution, and use of these genes.

Monomethylamine, Dimethylamine, Trimethylamine

The monomethylamine (MMA), dimethylamine (DMA), and trimethylamine (TMA) regions on the mtaC tree are relatively closely related and are often overlapping. This makes sense given that there is significant overlap in the proteins that facilitate conversion of methylamines to methanol, as well as the organisms that are able to grow on these substrates. As shown in Figure 3-2, which depicts the super pathway of methanogenesis, the MMA, DMA, and TMA pathways are unique in the corrinoid and MT1 proteins used (Table 3-1), but feed an intermediate to the same MT2. These pathways also differ in the early steps in the intermediate products that are released, ammonium for MMA, methylamine for DMA, and dimethylamine for TMA. This means that organisms capable of using all substrates would also be capable of using the products released by the early steps of the pathway (as in the case of DMA release) if they possess the proteins for multiple methylated substrates. These genes are indeed located in close proximity to one another. In *Methanosarcina barkeri*, the gene encoding the corrinoid protein, mtmC, is located directly upstream of the gene encoding MT1, mtmB. The gene encoding MT2 (mtbA), which is not specific to methylamine is located upstream of mtmC (Burke et al., 1998). Analogous proteins are present for DMA and TMA (Ferguson et al., 2000; Ferguson and Krzycki, 1997).

The MMA region of the tree includes the methanogenic Archaea *Methanosarcina*, *Methanolobus*, *Methanococcoides*, *Methanohalophilus*, *Methanomethylovorans*, and *Methanosalsum*. Apparent outgroups include Bacteria which possess putatively DMA-specific proteins including: *Desulfosporosinus*, *Eubacterium*, *Peptococcaceae*, *Desulfotomaculum*, and *Pseudobacteroides*. The DMA region of the tree includes the methanogenic Archaea *Methanosarcina*, *Methanomassiliicoccus*, *Methanococcoides*, *Methanohalophilus*, *Methanolobus*, *Methanosalsum*, and *Methanohalobium*. Apparent outgroups include Archaea with unidentified substrates and TMA-specific substrates. Using BLAST to search the NCBI database for some example taxa (e.g., WP_048107480.1) reveals many sequences annotated only as corrinoid protein, with just a few that are annotated as DMA. It is likely that sequences lacking substrate annotations in this region of the tree actually possess DMA or methylamine proteins. The TMA-specific sequences include *Methanosarcina*, *Methanococcoides*, *Methanophilus*, *Methanomethylovorans*, *Methanolobus*, and *Thermoplasmatales*. The apparent outgroup to the TMA region (and the three methylamine regions) include archaeal taxa including unidentified Euryarchaeotes, *Hadesarchaea*, and *Bathyarchaeota*.

The overall topology of the three methylamine-specific regions of the tree illustrates well-resolved substrate specificity, positioned within other bacterial taxa. This suggests these genes share an evolutionary history, possibly transferred from the phylogenetically-distant bacteria. The bacterial outgroup to the methylamine-specific sequences include *Sporomusa*,

Pelosinus, *Clostridium*, *Syntrophobotulus*, *Anaerovibrio*, and *Treponema*, taxa which may be assessed further for evolutionary links to methylotrophic methanogens.

Methanol

The methanol-specific region of the mtaC gene tree includes Archaea including *Methanosarcina*, *Methanlobus*, *Methanococcoides*, *Methanohalophilus*, *Methanomethylovorax*, *Methanosalsum*, *Methanoplasma*, *Hadesarchaea*, *Methanomassiliicoccus*, *Methanobrevibacter*, and *Methanosphaera*. It is notable that *Methanobrevibacter* (e.g., *M. smithii*), *Methanoplasma* (e.g., *Methanoplasma termitum*) *Methanosphaera* (e.g., *M. stadtmanae*) are present only in this substrate-specific region of the tree. In some ways, this serves as a positive control for validating the sequence hits on this gene tree in general. This is because *Methanobrevibacter* and *Methanosphaera* use a methanogenic pathway unique from all other methylotrophic methanogens in that they require the presence of hydrogen to use the substrate methanol. *Methanoplasma termitum* has also been shown to employ hydrogen-dependent methyltrophic methanogenesis (Lang et al., 2015). In general, *Methanobrevibacter* have some of the most limited metabolisms of all methanogenic Archaea. They are even unable to reduce carbon dioxide to methane (Fricke et al., 2006). Thus, it would be predicted that they would only possess methanol-specific sequences. A few Bacteria are also present in the methanol region including *Thermincola*, *Desulfosporosinus*, *Methanocella*, and *Desulfbulbus*. This may be indicative of a transfer to or from these Archaea, and should be evaluated in future work.

3.4.5 Taxon Sampling: Opportunities and Limitations

Taxon sampling is critical to gene and species tree construction, providing both insight and its own biases. The taxa included in the mtaC gene tree presented here yield some important findings, but also introduce areas for future work and improvement. Table 3-4 lists the taxa on the mtaC gene tree, in alphabetical order and with associated protein IDs and taxonomy. Of the all organisms present (852) most are archaea (756) and a smaller fraction are bacteria (96). Methanogenic archaea make up most of these archaeal species (>700), belonging to genera such as *Methanosarcina* (351), *Methanomassiliicoccus* (35), *Methanomethylophilus* (11), *Methanococcoides* (39), *Methanlobus* (49), and a few less well-represented groups. It is apparent from both the gene tree and taxa list that there are multiple copies of methyltransferases present in many methanogens. This indicates that methanogens are capable of employing multiple methylotrophic pathways using multiple substrates.

The distribution and topologies of many taxa on this tree are consistent with the literature. For instance, it is to be expected from previous work that Methanosarcinales have the substrate-specific proteins for all substrates, while *Methanosphaera* only have the proteins for methanol-specific methylotrophic methanogenesis. However, other taxa on this tree,

e.g., *Ca.* Bathyarchaeota, which appear to have the genetic information for multiple substrates, have conflicting reports of substrate utilization in the literature. Some studies report that *Ca.* Bathyarchaeota possess the genes capable of methylotrophic methanogenesis from methanol, methylamines, and methyl sulfides (Castelle and Banfield, 2018), while others report that they lack these genes (Vanwonterghem et al., 2016). It is clear from this analysis that Bathyarchaeota do have the genetic capability for methylotrophic methanogenesis, but it remains to be seen what they do in culture. Other taxa, such as *Ca.* 'Korarchaeum cryptofilum' (WP_012309677.1), which is an archaeon grouped within the largest group of bacteria on the mtaC tree, may provide clues for understanding the evolution of archaea at large (Elkins et al., 2008).

Methylotrophic methanogenesis is arguably the least well-understood methanogenic pathway, perhaps because the phylogenetic distribution of genes that encode this pathway are complex, were found later, or because the organisms that are capable of this pathway are often found in some of the most remote environments (and thus, it is likely that additional taxa remain to be sampled or cultured). Many studies also note the need and value of studying methyl-corrinoid proteins such as mtaC in particular in the future (Kolb and Stacheter, 2013; Sousa et al., 2018). Therefore, it is an exciting time to further understand the phylogenetic distribution of methylotrophic genes both in methanogenic lineages as well as some of their nearest (phylogenetic and geographic) neighbors.

See this text in Appendix F.

3.4.6 Interrogating Microbial Carbon Metabolism in Marine Sediments with Methyl-Corrinoids

Methylotrophic methanogens are common in marine sediments, and the number of new putatively methanogenic taxa that possess the genes for this metabolism is rapidly growing (Lyu et al., 2018). As is noted throughout this discussion, many of the organisms that carry out methylotrophic methanogenesis or harbor a methyl corrinoid protein are associated with the marine environment. Methane cycling is a poorly understood but important metabolic process in marine sedimentary environments (Ijiri et al., 2018; Marlow et al., 2016; Orcutt et al., 2011; Valentine, 2011). The largest reservoir of methane on Earth is found in marine sediments (Orcutt et al., 2011). Moreover, methylotrophic methanogenesis is likely a key metabolism for better understanding carbon and sulfur cycling dynamics in the marine environment (Summons et al., 1998). Although there are many areas of uncertainty in understanding the carbon cycle in marine sediments, methanol and methyl compounds are increasingly noted as important substrates for microbes in deep-subsurface environments, particularly because in sulfate-rich environments, sulfate reducers cannot outcompete methylotrophic methanogens for certain substrates (Yanagawa et al., 2016). Nonetheless, how sulfate-reducing bacteria employ methyltransferases is not entirely clear (Sousa et al.,

2018). Interestingly, *Desulfotomaculum* is one bacterium that appears on this tree. It has been shown that this organism lives in deep, subsurface sediments, and possesses the machinery to utilize two different methanol-degradation pathways (Sousa et al., 2018).

Many bacteria and archaea on this tree have connections to oligotrophic or subsurface marine systems. For instance, *Candidatus* 'Aminicenantes' (OGD11856.1) which appears in the largest bacterial clade, is a newly described bacterial lineage distributed throughout marine environments such as hydrothermal vents and coral-associated microbiomes (Farag et al., 2014). *Nitrosopumilus maritimus* (ABX13163.1), an Archaeon that is present in the methionine region of the tree has been suggested to play an important role in carbon and nitrogen cycling in marine environments (Könneke et al., 2005; Walker et al., 2010; Bayer et al., 2016). The ammonia-oxidizing Archaeon *Candidatus* 'Nitrosopelagicus brevis' (WP_048105639.1) also present on the methionine region of the tree, was also isolated from the oligotrophic ocean (Santoro et al., 2015). *Methanotorris igneus* (WP_013798831.1), a methanogen isolated from a black smoker, was shown not to have enhanced growth in the presence of methylated substrates, but clearly possesses these sequences (Takai et al., 2004). Because this particular taxon sits near the Bacteria as an unidentified substrate, it is possible that the capabilities of this methanogen should be further assessed. *Staphylothermus marinus* (WP_011838717.1) was isolated from a hydrothermal vent on the East Pacific Rise (Anderson et al., 2009). *Thermosediminibacter oceani* (WP_013275879.1) was isolated from deep sea sediments of Peru Margin (Lee et al., 2005). *Parasporobacterium paucivorans* (WP_073994062.1) is a Bacterium which groups with the dimethylamine region of the tree. It was isolated from a freshwater environment, but actually appears to produce DMS and methanethiol (Lomans et al., 2001) suggesting a possible microbial source for these compounds in the environment. The production of compounds such as methanethiols in the environments from which these microbes have been isolated has only been recently measured, in large part, due to improved methods for studying, measuring, and culturing these systems (Reeves et al., 2012, 2014; McNichol et al., 2016). It has been shown that methanethiols can form in hydrothermal vent environments either abiotically or from the thermal degradation of organic matter which may contribute to the metabolic strategies employed by microbes in these systems (Reeves et al., 2014). The origin of these substrates must be further assessed in terms of the genes that act upon them to better understand the microbial metabolic dynamics of marine sedimentary environments. Nonetheless, methyl-corrinoids proteins may provide an useful starting point.

3.4.7 Outlook: Evaluating DMS Metabolism and the Timing of Methyl Specificity in Methanogenic and Bacterial Lineages

The metabolism of methylated thiols such as DMS may have even greater importance in marine sediments than previously realized. Further analysis of DMS metabolism genes of

interest (Table 3-2) and producers of DMS in the environment (*e.g.*, Table 3-3) should be assessed to disentangle possible relationships between the evolutionary histories of DMS-producing organisms. In particular, timing the emergence of the producers of methylated thiols and other methyl substrates as well as the consumers will be an exciting next step to understand how metabolisms have evolved and affected the biogeochemistry over the last two eons. Nonetheless, this will require a more extensive assessment of methylotrophic gene histories.

The mtaC tree topology presented here shows clear substrate specificity, but not all regions are well supported. Future work may consider applying alignment techniques such as a profile alignment or concatenation of methyltransferase MT1, MT2, and corrinoid proteins together. This may aid in increasing tree support and more clearly defining the root of methanogenic methyl corrinoid proteins in the mtaC tree which would greatly improve inferences surrounding methanogen evolution.

Another exciting approach for future work is the possibility of dating substrate specificity in methylotrophic lineages by tethering the temporal information provided by bacterial lineages such as Firmicutes (as discussed above) or members of Actinobacteria, Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, and Ascomycota which are hypothesized to also harbor methyl corrinoid genes (Kolb and Stacheter, 2013).

3.5 Conclusions

In summary, we present the findings of our work to further illuminate the phylogenetic history of methyl corrinoid proteins, environmentally significant enzymes in the biogeochemical cycling of carbon. The phylogenetic history of methyl-corrinoid genes is more complex than first hypothesized; there is not a clear pattern between bacterial and archaeal lineages and it appears that there may have been multiple transfers or duplications. Bacteria seem to be implicated in these events, but polarizing these transfers and dating the timing of substrate specificity was not possible with this dataset. Many of the taxa involved in this analysis are related to the marine environment, but it is not yet possible to conclude any environmental link between gene history or substrate specificity. Nonetheless, there is a clear pattern of substrate specificity within archaeal taxa, which may still hold promise for disentangling a phylogenetic history of methanogens.

The methyl corrinoid pathway of methanogenesis may provide novel insights into the metabolic capabilities of recently described taxa, and may also reveal important links in the biogeochemical cycling of elements such as carbon, nitrogen, and sulfur in the marine system. It seems plausible that many of these sequences were distributed via HGT from bacteria, perhaps even in the marine environment. Emerging methods in molecular evolution will be useful, particularly for interrogating the substrate specificity of the methylotrophic

pathway. Combined with exploration into environments such as ocean sediments or the deep subsurface, molecular evolution analyses do hold promise for better understanding the biogeochemistry and methane dynamics of the Earth's most globally-significant reservoirs of methane.



Figure 3-1: Timeline of events in the evolution of the carbon cycle highlighting events relevant to methanogenesis, shifting from a dominantly abiogenic methane world to a dominantly microbial methane world. The 4.5 Ga time period is set by zircons (Wilde et al., 2001), which suggest that the Earth had liquid water at 4.4 Ga. The origin and diversification of life remains an open question, but is taken to be sometime before our earliest evidence before the earliest "core" pathway of microbial methanogenesis. The first methanogenic pathway illustrated here is defined by fluid inclusions (Ueno et al., 2006), isotopically depleted kerogen (Hinrichs, 2002), and molecular clock estimates (Wolfe and Fournier, 2018). The methylotrophic pathway of methanogenesis is labeled with an unknown time. The acetoclastic pathway of methanogenesis occurs around 250 Ma based on phylogenetic estimates (Rothman et al., 2014).

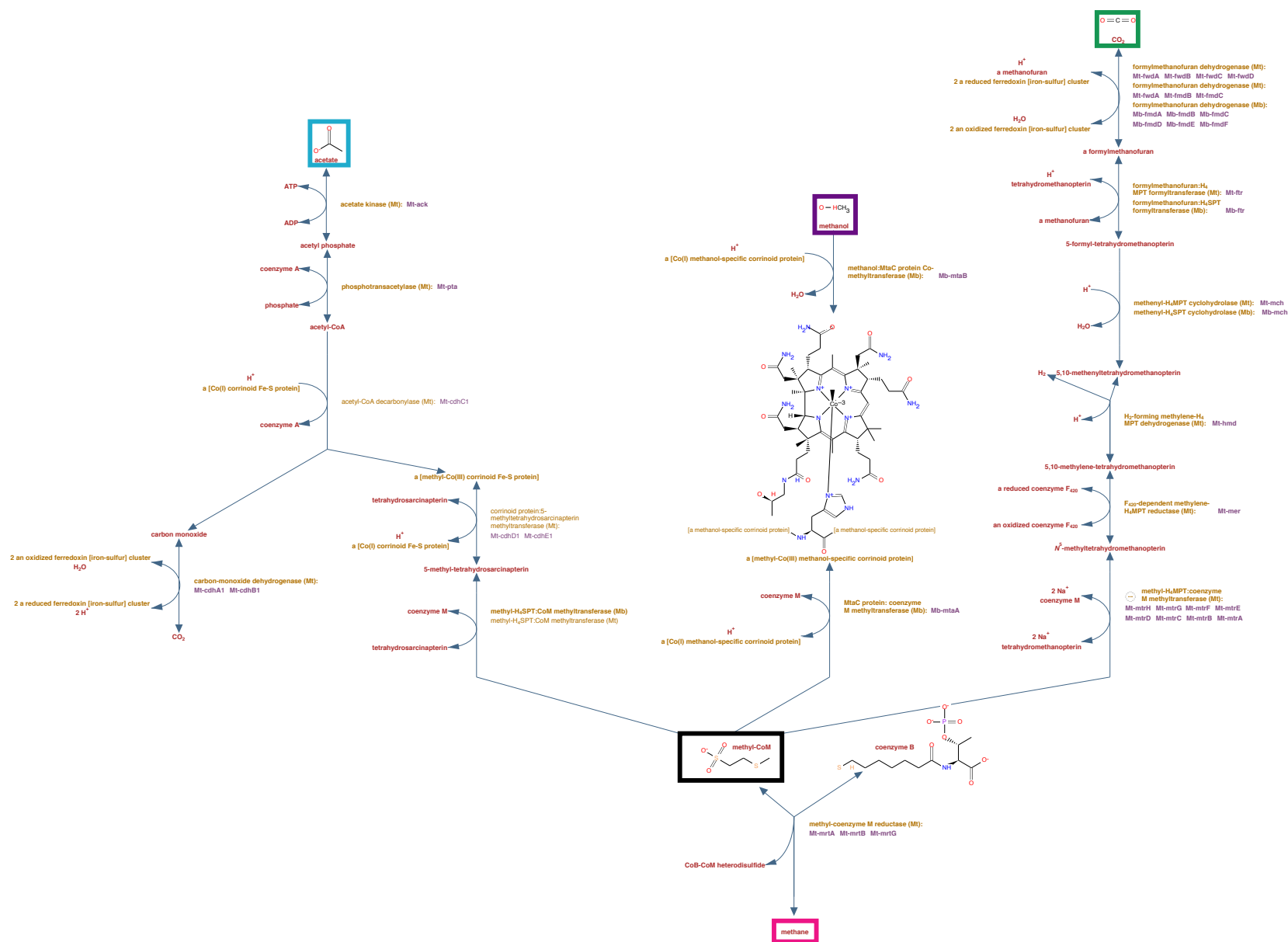


Figure 3-2: Methanogenesis Pathways. The underlying pathways were generated by MetaCyc, but were modified for readability.

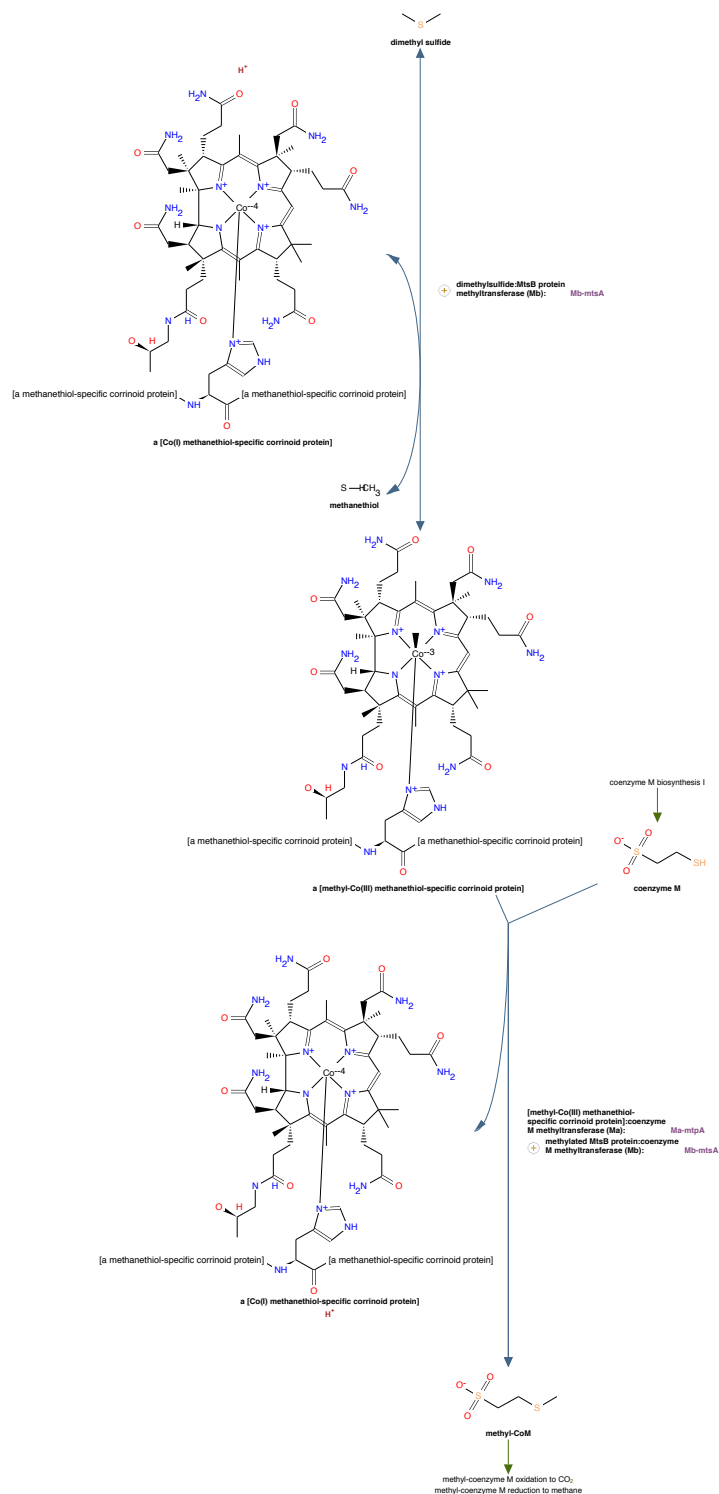


Figure 3-3: Pathway of Methylotrophic Methanogenesis (from DMS). Modified from the pathway diagram generated by MetaCyc (Caspi et al., 2014) and including enzyme information from *Methanosarcina barkeri* in the methyl-corrinoid pathway, specifically for utilization of the substrate DMS.



Figure 3-4: Simplified Methyl-Corrinoid Pathway. The corrinoid proteins are in gold, and the MT1 and MT2 labels indicate the methyltransferase (1 and 2) steps. The corrinoid protein and MT1 are considered to be substrate specific, while MT2 is not.

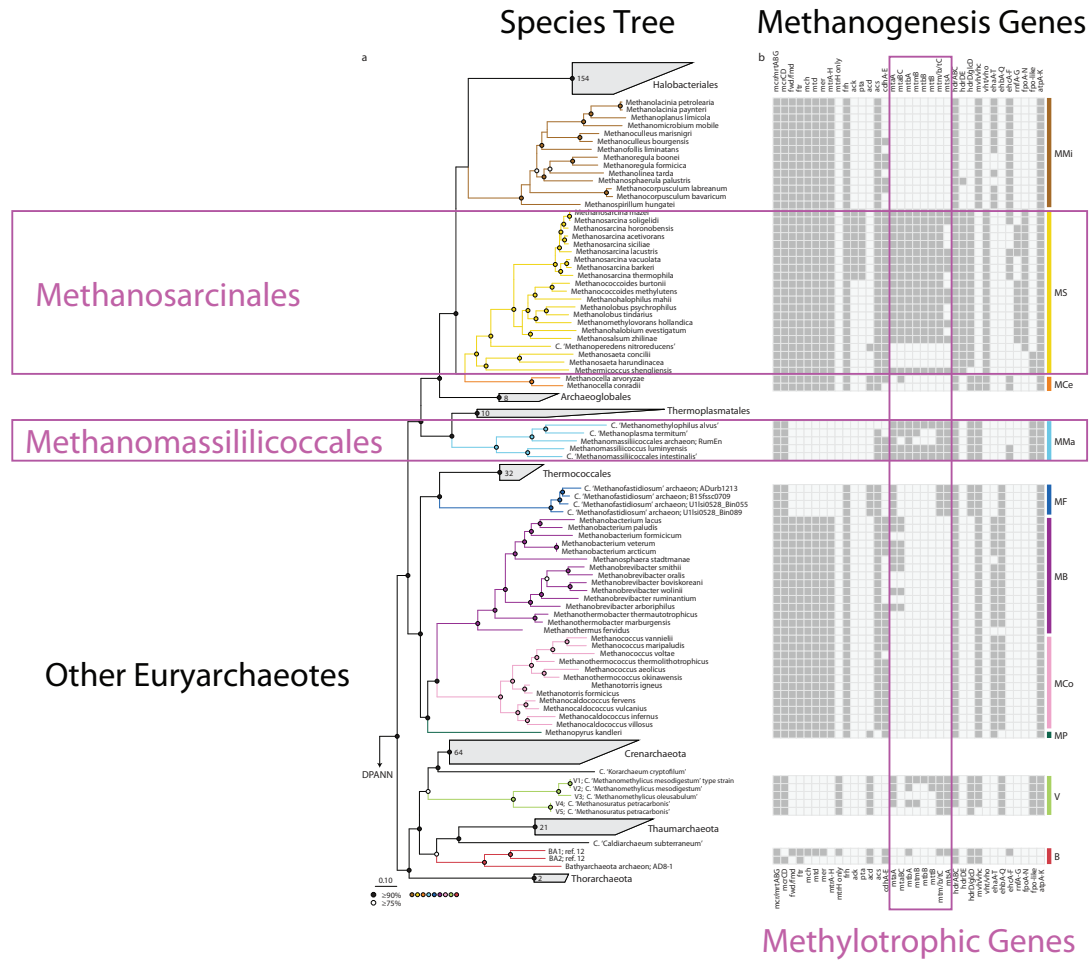


Figure 3-5: A presence-absence diagram depicting both the phylogenetic relationships between taxa on the species tree and the genes associated with each pathway modified from Vanwonterghem et al. (2016). The genes associated with methylotrophic methanogenesis are highlighted.

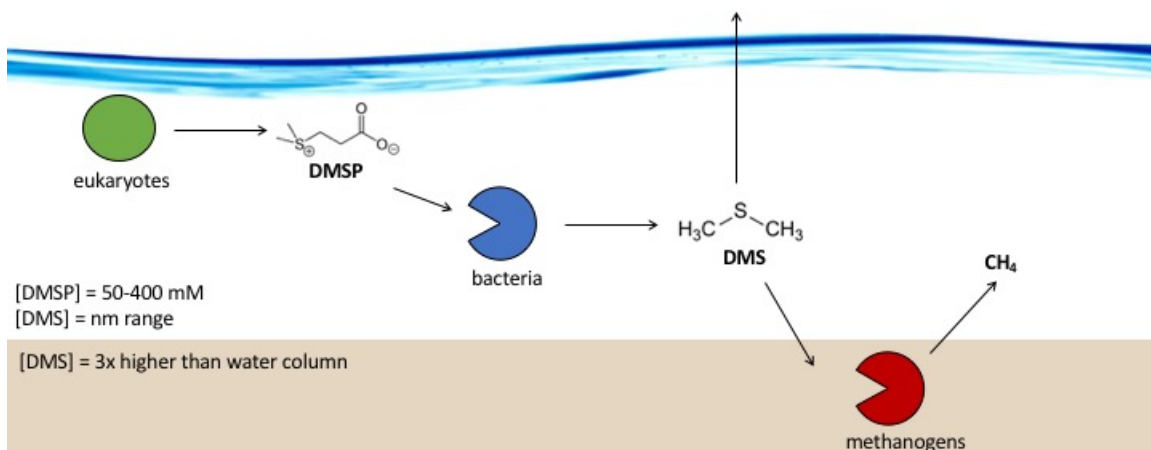


Figure 3-6: Relationship between major mediators of DMSP production and degradation in the marine environment. This diagram includes only the major producers and consumers of each molecule for simplification (Agalidis et al., 1997; Munk et al., 2011). Values of DMSP and DMS represent estimates.

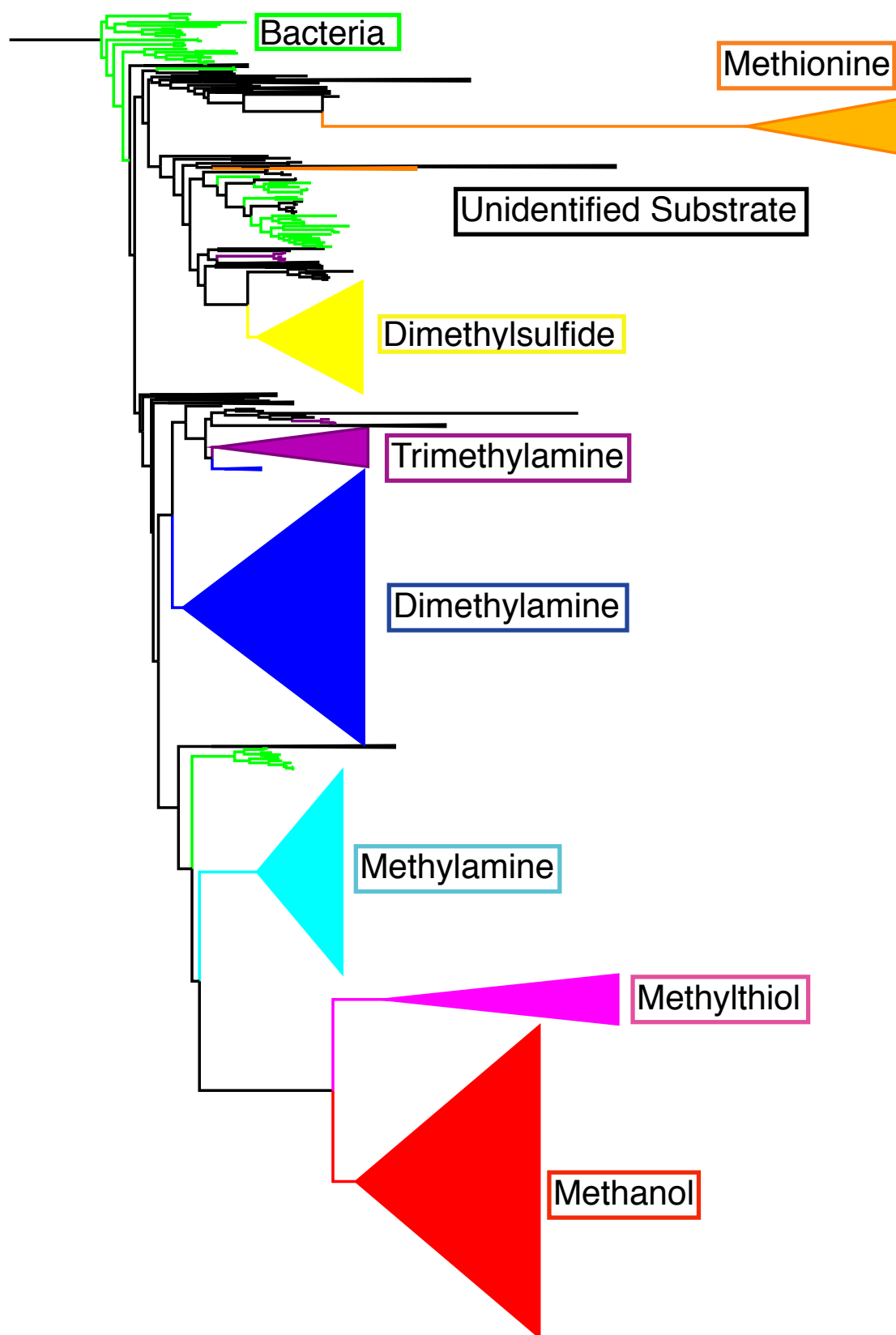


Figure 3-7: Gene tree of methyl-corrinoid (mtaC) homologs, colored according to NCBI substrate annotation and manual review of sequences. This tree was built with RAxML and rooted within bacteria by manual assessment of taxa as described in methods.

Table 3-1: Methyltransferase pathway genes of interest. Information derived from MetaCyc and literature (Borrel et al., 2014; Curson et al., 2011; Galagan et al., 2002; Vanwonterghem et al., 2016).

Gene	Protein Name	Substrate	Function
mtbB1	Dimethylamine methyltransferase	dimethylamine	MT1
mtbA	Methylcobamide:CoM methyltransferase MtbA	dimethylamine	MT2
mtbC	Dimethylamine corrinoid protein	dimethylamine	corrinoid
mtsA	Methylated-thiol-coenzyme M methyltransferase	dimethylsulfide	MT1
mtpA or cmtA	Methylcobalamin:CoM methyltransferase isozyme A	dimethylsulfide	MT2
mtsB	Methylated-thiol-corrinoid protein	dimethylsulfide	corrinoid
mtsA	Methylated-thiol-coenzyme M methyltransferase	methanethiol	MT1
mtpA or cmtA	Methylcobalamin:CoM methyltransferase isozyme A	methanethiol	MT2
MtsB	Methylated-thiol-corrinoid protein	methanethiol	corrinoid
mtaB	Methanol-corrinoid protein co-methyltransferase	methanol	MT1
mtaA	Methylcobamide:CoM methyltransferase MtaA	methanol	MT2
mtaC	Methanol-corrinoid protein	methanol	corrinoid
mtmB1	Monomethylamine methyltransferase	methylamine	MT1
mtbA	Methylcobamide:CoM methyltransferase	methylamine	MT2
mtmC1	Monomethylamine corrinoid protein 1	methylamine	corrinoid
mtsA	Methylated-thiol-coenzyme M methyltransferase	methylthiopropionate	MT1
mtpA or cmtA	Methylcobalamin:CoM methyltransferase isozyme A	methylthiopropionate	MT2
MtsB	Methylated-thiol-corrinoid protein	methylthiopropionate	corrinoid
MtqB	tetramethylammonium methyltransferase	teramethylammonium	MT1
MtqA	methylcobamide:CoM methyltransferase	teramethylammonium	MT2
MtqC		teramethylammonium	corrinoid
mttB	Trimethylamine methyltransferase	trimethylamine	MT1
mtbA	methylcobamide:CoM methyltransferase	trimethylamine	MT2
mttC	Trimethylamine corrinoid protein	trimethylamine	corrinoid

Table 3-2: DMS Genes of Interest.

Pathway	Gene	Enzymatic Activity	Organism
DMSP to DMS	dddP	dimethylsulfoniopropionate lyase	Mainly in roseobacters, but also in 'Candidatus Puniceispirillum marinum str. IMCC1322' (in the SAR116 clade), <i>Oceanimonas doudoroffii</i> (a gammaproteobacterium) and some ascomycete fungi
DMSP to DMS	dddD	DMSP cleavage to DMS	Sporadic occurrence in alphaproteobacteria, betaproteobacteria and, more frequently, gammaproteobacteria; often found in bacteria that grow well on DMSP as the sole carbon source
DMSP to DMS	dddL	DMSP lysis	In marine alphaproteobacteria, mainly roseobacters
DMSP to DMS	dddQ	DMSP lysis	Exclusively in roseobacters
DMSP to DMS	dddW	DMSP lysis	Exclusively in roseobacters, but only in two strains
DMSP to DMS	dddY	DMSP lysis	Sporadic occurrence in betaproteobacteria, gammaproteobacteria, deltaproteobacteria and epsilonproteobacteria
DMSP to DMS	Alma1	dimethylsulfoniopropionate lyase	<i>Emiliana huxleyi</i> (bloom forming algae)
DMSP to DMS	dmdA	DMSP demethylation	SAR11 bacteria and roseobacters

Table 3-3: DMSP sources, relevant organisms, and proposed evolutionary dates.

Taxa	Clade	Date..approx..if.known.	Reference
Hymenomonas carterae	Coccolithophore		(Vairavamurthy et al., 1985)
Gymnodinium nelsoni	Dinoflagellate	expansion Jurassic	(Dacey and Wakeham, 1986)
Platymonas subcordiformis	Prasinophytes		(Dickson and Kirst, 1986)
Phaeocystis sp.			(Stefels and Boekel, 1993)
Melosira numuloides (Diatoms)		100 Ma	(Keller, 1989)
Chrysamoeba sp. (chrysophyte)			(Yoch, 2002)
Ochromonas sp. (chrysophyte)			(Yoch, 2002)
Prorocentrum sp. strain IIB2b1 (dinoflagellate)		expansion Jurassic	(Yoch, 2002)
Emiliana huxleyi BT6 (coccolithophore)	Haptophyte		(Yoch, 2002)
Dinoflagellates		expansion Jurassic	
Dinoflagellates		>250 Ma	(Parfrey et al., 2011)
Diatoms et al (Stramenopiles)		>750 Ma	(Parfrey et al., 2011)
Haptophytes		>575 Ma	(Parfrey et al., 2011)

Table 3-4: Taxa listed in Figure 3-7.

Taxon Name	Protein Description	Protein ID	Taxonomy
Acetothermia_bacterium_64_32	Cobalamin	KUK27705.1	Bacteria
Acidobacteria_bacterium_13_1_20CM_2_68_14	hypothetical	OLD65654.1	Bacteria
Acidobacteria_bacterium_13_1_20CM_2_68_7	hypothetical	OLE67755.1	Bacteria
Acidobacteria_bacterium_13_1_40CM_2_68_5	hypothetical	OLD63858.1	Bacteria
Actinobacteria_bacterium_RBG_19FT_COMBO_70_19	hypothetical	OFW77058.1	Bacteria
Anaerolineae_bacterium_SM23_63	hypothetical	KPK90936.1	Bacteria
Anaerovibrio_lipolyticus_DSM_3074	cobalamin-binding	WP_039205959.1	Bacteria
Anaerovibrio_lipolyticus	cobalamin-binding	WP_027397382.1	Bacteria
Anaerovibrio_sp_RM50	cobalamin-binding	WP_027407411.1	Bacteria
Arc_I_group_archaeon_ADurb1113_Bin01801	Trimethylamine	KYC55042.1	Archaea
Arc_I_group_archaeon_ADurb1213_Bin02801	Methylated-thiol-coenzyme	KYC55383.1	Archaea
Arc_I_group_archaeon_ADurb1213_Bin02801	Trimethylamine	KYC58006.1	Archaea
Arc_I_group_archaeon_BMIXfssc0709_Meth_Bin006	Methylated-thiol-coenzyme	KYC45501.1	Archaea
Arc_I_group_archaeon_BMIXfssc0709_Meth_Bin006	Trimethylamine	KYC44725.1	Archaea
Arc_I_group_archaeon_U1lsi0528_Bin055	Methanol-corrinoid	KYC48099.1	Archaea
Arc_I_group_archaeon_U1lsi0528_Bin055	Trimethylamine	KYC50875.1	Archaea
Arc_I_group_archaeon_U1lsi0528_Bin089	Trimethylamine	KYC51555.1	Archaea
Archaeoglobus_fulgidus_DSM_8774	cobalamin-binding	WP_010877520.1	Archaea
Archaeoglobus_fulgidus_DSM_8774	corrinoid	AAB91218.1	Archaea
Archaeoglobus_fulgidus	cobalamin-binding	WP_048064581.1	Archaea
archaeon_Heimdall_LC_3	Trimethylamine	OLS22613.1	Archaea
candidate_division_Zixibacteria_bacterium_SM23_81	methyltransferase	KPL19208.1	Archaea
candidate_division_MSBL1_archaeon_SCGC-AAA259D18	hypothetical	KXA90803.1	Archaea
candidate_division_MSBL1_archaeon_SCGC-AAA259D18	hypothetical	KXA91218.1	Archaea
candidate_division_MSBL1_archaeon_SCGC-AAA259E17	methyltransferase_	KXA92702.1	Archaea
candidate_division_MSBL1_archaeon_SCGC-AAA259E19	hypothetical	KXA94216.1	Archaea
candidate_division_MSBL1_archaeon_SCGC-AAA259E19	hypothetical	KXA94469.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
candidate_divison_MSBL1_archaeon_SCGC-AAA259E19	hypothetical	KXA95296.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA259M10	hypothetical	KXA99102.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA259O05	hypothetical	KXB00676.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA261F19	cobalamin-binding	KXB03897.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA382A03	hypothetical	KXB04090.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA382A20	cobalamin-binding	KXB06230.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA382A20	hypothetical	KXB06477.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA382A20	hypothetical	KXB07048.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA382A20	hypothetical	KXB07590.1	Archaea
candidate_divison_MSBL1_archaeon_SCGC-AAA382C18	methanol_corrinoid	KXB06533.1	Archaea
Candidatus_Altiarchaeales_archaeon_WOR_SM1_79	hypothetical	ODS36526.1	Bacteria
Candidatus_Aminicenantes_bacterium_RBG_13_62_12	dimethylamine	OGD12780.1	Bacteria
Candidatus_Aminicenantes_bacterium_RBG_16_63_16	dimethylamine	OGD22292.1	Bacteria
Candidatus_Aminicenantes_bacterium_RBG_16_66_30	dimethylamine	OGD11856.1	Bacteria
Candidatus_Aminicenantes_bacterium_RBG_19FT_COMBO_58_17	dimethylamine	OGD37535.1	Bacteria
Candidatus_Aminicenantes_bacterium_RBG_19FT_COMBO_59_29	dimethylamine	OGD26035.1	Bacteria
Candidatus_Bathyarchaeota_archaeon_B23	Trimethylamine	KYH39426.1	Archaea
Candidatus_Bathyarchaeota_archaeon_B24	corrinoid	KYH37114.1	Archaea
Candidatus_Bathyarchaeota_archaeon_B26-1	corrinoid	KYH42497.1	Archaea
Candidatus_Bathyarchaeota_archaeon_B26-2	corrinoid	KYH40176.1	Archaea
Candidatus_Bathyarchaeota_archaeon_B63	putative	KYH40379.1	Archaea
Candidatus_Bathyarchaeota_archaeon_BA1	Dimethylamine	KPV61612.1	Archaea
Candidatus_Bathyarchaeota_archaeon_BA1	Dimethylamine	KPV64356.1	Archaea
Candidatus_Bathyarchaeota_archaeon_BA1	Dimethylamine	KPV64367.1	Archaea
Candidatus_Bathyarchaeota_archaeon_BA1	Dimethylamine	KPV64966.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_13_38_9	hypothetical	OGD53017.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_13_38_9	hypothetical	OGD53045.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_13_38_9	hypothetical	OGD54883.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Candidatus_Bathyarchaeota_archaeon_RBG_13_38_9	hypothetical	OGD54910.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_13_38_9	hypothetical	OGD54960.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_13_52_12	hypothetical	OGD60770.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_16_48_13	hypothetical	OGD46258.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_16_48_13	hypothetical	OGD46569.1	Archaea
Candidatus_Bathyarchaeota_archaeon_RBG_16_57_9	hypothetical	OGD45974.1	Archaea
Candidatus_Korarchaeum_cryptofilum_OPF8	corrinoid	WP_012309677.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	cobalamin-binding	WP_020448778.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	dimethylamine	WP_020448787.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	Methanol	AGN25680.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	Methanol	AGY50164.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	methanol	WP_020448235.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	methionine	WP_048134245.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	trimethylamine	WP_020448772.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	trimethylamine	WP_020448782.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis_Issoire-Mx1	trimethylamine	WP_020449409.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis	methanol–corrinoid	WP_048133854.1	Archaea
Candidatus_Methanomassiliicoccus_intestinalis	methanol–corrinoid	WP_048134198.1	Archaea
Candidatus_Methanomethylophilus_alvus_Mx1201	5-methyltetrahydrofolate–homocysteine	WP_015504410.1	Archaea
Candidatus_Methanomethylophilus_alvus_Mx1201	dimethylamine	WP_015505018.1	Archaea
Candidatus_Methanomethylophilus_alvus_Mx1201	hypothetical	WP_015505015.1	Archaea
Candidatus_Methanomethylophilus_alvus_Mx1201	methionine	WP_015504315.1	Archaea
Candidatus_Methanomethylophilus_alvus_Mx1201	methylthiol–coenzyme	WP_015505223.1	Archaea
Candidatus_Methanomethylophilus_alvus_Mx1201	trimethylamine	WP_015504327.1	Archaea
Candidatus_Methanomethylophilus_sp_1R26	dimethylamine	WP_058746969.1	Archaea
Candidatus_Methanomethylophilus_sp_1R26	dimethylamine	WP_058747236.1	Archaea
Candidatus_Methanomethylophilus_sp_1R26	methanol–corrinoid	WP_058747807.1	Archaea
Candidatus_Methanomethylophilus_sp_1R26	methionine	WP_058748236.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Candidatus_Methanomethylophilus_sp_1R26	methylthiol-CoM	WP_058747653.1	Archaea
Candidatus_Methanoplasma_termitum	dimethylamine	WP_048111918.1	Archaea
Candidatus_Methanoplasma_termitum	methanol-corrinoid	WP_048111798.1	Archaea
Candidatus_Methanoplasma_termitum	methionine	WP_052399287.1	Archaea
Candidatus_Methanoplasma_termitum	putative	AIZ56770.1	Archaea
Candidatus_Nitrosoarchaeum_koreensis_MY1	methionine	WP_007551119.1	Archaea
Candidatus_Nitrosoarchaeum_limnia_BG20	B12	EPA05056.1	Archaea
Candidatus_Nitrosoarchaeum_limnia_SFB1	methionine	EGG42114.1	Archaea
Candidatus_Nitrosopelagicus_brevis	methionine	WP_048105639.1	Archaea
Candidatus_Nitrosopumilus_adriaticus	methionine	WP_048115673.1	Archaea
Candidatus_Nitrosopumilus_koreensis_AR1	MULTISPECIES_	WP_014963694.1	Archaea
Candidatus_Nitrosopumilus_piranensis	methionine	AJM92058.1	Archaea
Candidatus_Nitrosopumilus_salaria_BD31	methionine	WP_008301613.1	Archaea
Candidatus_Nitrosopumilus_sp_AR2	MULTISPECIES_	WP_014965586.1	Archaea
Candidatus_Nitrososphaera_evergladensis_SR1	methionine	AIF85021.1	Archaea
Candidatus_Nitrososphaera_gargensis_Ga9_2	putative	AFU57439.1	Archaea
Candidatus_Nitrosotalea_devanattera	Methionine	CUR52375.1	Archaea
Candidatus_Nitrosotenuis_cloacae	methionine	WP_048187301.1	Archaea
Candidatus_Syntrophoarchaeum_butanivorans	corrinoid	OFV65992.1	Archaea
Candidatus_Syntrophoarchaeum_butanivorans	corrinoid	OFV66410.1	Archaea
Candidatus_Syntrophoarchaeum_caldarius	corrinoid	OFV67276.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ-45	hypothetical	KXH70679.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ-45	hypothetical	KXH72076.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ1-45	hypothetical	KXH70073.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ1-45	hypothetical	KXH70541.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ1-45	hypothetical	KXH74033.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ1-83	hypothetical	KXH70129.1	Archaea
Candidatus_Thorarchaeota_archaeon_SMTZ1-83	hypothetical	KXH74024.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Candidatus_Thorarchaeota_archaeon_SMTZ1-83	hypothetical	KXH77722.1	Archaea
Cenarchaeum_symbiosum_A	methionine	ABK77567.1	Archaea
Chloroflexi_bacterium_GWC2_73_18	methyltransferase	OGN88425.1	Bacteria
Chloroflexi_bacterium_RBG_13_50_21	hypothetical	OGN97568.1	Bacteria
Chloroflexi_bacterium_RBG_16_47_49	hypothetical	OGO12010.1	Bacteria
Chloroflexi_bacterium_RBG_16_52_11	methyltransferase	OGO25865.1	Bacteria
Chloroflexi_bacterium_RBG_16_54_11	hypothetical	OGO27286.1	Bacteria
Chloroflexi_bacterium_RBG_16_54_18	hypothetical	OGO33135.1	Bacteria
Clostridium_botulinum_CDC_1436	cobalamin-binding	WP_012720395.1	Bacteria
Clostridium_botulinum	cobalamin-binding	WP_075860993.1	Bacteria
Clostridium_carboxidivorans_P7	cobalamin-binding	WP_007060633.1	Bacteria
Clostridium_cellulovorans_743B	cobalamin-binding	WP_010076138.1	Bacteria
Clostridium_ljungdahlii	cobalamin-binding	WP_063554956.1	Bacteria
Clostridium_purinilyticum	cobalamin-binding	WP_050356052.1	Bacteria
Clostridium_scatologenes	MULTISPECIES_	WP_029160050.1	Bacteria
Clostridium_senegalense	cobalamin-binding	WP_010292018.1	Bacteria
Clostridium_sp_KNHs214	cobalamin-binding	WP_035294784.1	Bacteria
Clostridium_straminisolvans_JCM_21531	cobalamin-binding	WP_038286537.1	Bacteria
Clostridium_tyrobutyricum	hypothetical	WP_017751388.1	Bacteria
Deltaproteobacteria_bacterium_CG2_30_43_15	cobalamin-binding	OIP31517.1	Bacteria
Deltaproteobacteria_bacterium_RBG_16_48_10	hypothetical	OGP92226.1	Bacteria
Deltaproteobacteria_bacterium_RBG_19FT_COMBO_46_12	hypothetical	OGP98518.1	Bacteria
Desulfitibacter_sp_BRH_c19	methyltransferase	KUO50702.1	Bacteria
Desulfitibacter_sp_BRH_c19	methyltransferase	KUO50706.1	Bacteria
Desulfitobacterium_chlororespirans_DSM_11544	trimethylamine	SHN85887.1	Bacteria
Desulfitobacterium_chlororespirans	dimethylamine	WP_072774686.1	Bacteria
Desulfitobacterium_hafniense	dimethylamine	WP_015945169.1	Bacteria
Desulfitobacterium_hafniense	methyltransferase	WP_035213634.1	Bacteria

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Desulfitobacterium_hafniense	methyltransferase	WP_058491970.1	Bacteria
Desulfobulbaceae_bacterium_BRH_c16a	methanol-5-hydroxybenzimidazolylcobamide	KJR96643.1	Bacteria
Desulfobulbus_japonicus	methanol-5-hydroxybenzimidazolylcobamide	WP_028582363.1	Bacteria
Desulfosporosinus_hippei_DSM_8344	trimethylamine	SDI34077.1	Bacteria
Desulfosporosinus_lacus_DSM_15449	dimethylamine	WP_073033384.1	Bacteria
Desulfosporosinus_meridiei_DSM_13257	corrinoid	WP_014903195.1	Bacteria
Desulfosporosinus_meridiei_DSM_13257	corrinoid	WP_014904786.1	Bacteria
Desulfosporosinus_meridiei_DSM_13257	dimethylamine	WP_014903796.1	Bacteria
Desulfosporosinus_orientis_DSM_765	cobalamin-binding	WP_014184027.1	Bacteria
Desulfosporosinus_orientis_DSM_765	corrinoid	WP_014186397.1	Bacteria
Desulfosporosinus_orientis_DSM_765	corrinoid	WP_014186640.1	Bacteria
Desulfosporosinus_orientis_DSM_765	corrinoid	WP_014187249.1	Bacteria
Desulfosporosinus_sp_BG	dimethylamine	WP_068966346.1	Bacteria
Desulfosporosinus_sp_BICA1-9	methanol-5-hydroxybenzimidazolylcobamide	KJS47453.1	Bacteria
Desulfosporosinus_sp_BICA1-9	methanol-5-hydroxybenzimidazolylcobamide	KJS89063.1	Bacteria
Desulfosporosinus_sp_BICA1-9	methyltransferase	KJS47628.1	Bacteria
Desulfosporosinus_sp_HMP52	dimethylamine	WP_034599966.1	Bacteria
Desulfosporosinus_sp_HMP52	methyltransferase	WP_034601983.1	Bacteria
Desulfosporosinus_sp_I2	methyltransferase	WP_045572706.1	Bacteria
Desulfosporosinus_sp_OL	cobalamin-binding	WP_075364787.1	Bacteria
Desulfosporosinus_sp_OL	Dimethylamine	OLN32034.1	Bacteria
Desulfosporosinus_sp_OT	cobalamin-binding	WP_040411893.1	Bacteria
Desulfosporosinus_sp_OT	dimethylamine	EGW39498.1	Bacteria
Desulfosporosinus_youngiae_DSM_17734	corrinoid	WP_007784993.1	Bacteria
Desulfosporosinus_youngiae_DSM_17734	corrinoid	WP_007785983.1	Bacteria
Desulfosporosinus_youngiae_DSM_17734	corrinoid	WP_007786830.1	Bacteria
Desulfotomaculum_acetoxidans_DSM_771	dimethylamine	WP_012813440.1	Bacteria
Eubacterium_aggregans	5-methyltetrahydrofolate–homocysteine	SEA30191.1	Bacteria

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Eubacterium_callanderi	5-methyltetrahydrofolate-homocysteine	SFP58394.1	Bacteria
Eubacterium_limosum	dimethylamine	WP_038352117.1	Bacteria
Euryarchaeota_archaeon_55_53	dimethylamine	WP_042686903.1	Archaea
Euryarchaeota_archaeon_55_53	Trimethylamine	KUK04061.1	Archaea
Euryarchaeota_archaeon_RBG_13_57_23	hypothetical	OGS43022.1	Archaea
Euryarchaeota_archaeon_RBG_13_57_23	hypothetical	OGS44410.1	Archaea
Euryarchaeota_archaeon_RBG_13_57_23	hypothetical	OGS44593.1	Archaea
Euryarchaeota_archaeon_RBG_13_61_15	hypothetical	OGS52213.1	Archaea
Euryarchaeota_archaeon_RBG_13_61_15	hypothetical	OGS52581.1	Archaea
Euryarchaeota_archaeon_RBG_16_62_10	hypothetical	OGS41508.1	Archaea
Euryarchaeota_archaeon_RBG_16_62_10	hypothetical	OGS41863.1	Archaea
Euryarchaeota_archaeon_RBG_16_67_27	hypothetical	OGS47549.1	Archaea
Euryarchaeota_archaeon_RBG_19FT_COMBO_56_21	hypothetical	OGS56899.1	Archaea
Euryarchaeota_archaeon_RBG_19FT_COMBO_56_21	hypothetical	OGS56929.1	Archaea
Euryarchaeota_archaeon_RBG_19FT_COMBO_69_17	hypothetical	OGS59937.1	Archaea
Euryarchaeota_archaeon_RBG_19FT_COMBO_69_17	hypothetical	OGS61399.1	Archaea
Ferroglobus_placidus_DSM_10642	cobalamin-binding	WP_012966230.1	Archaea
Hadesarchaea_archaeon_DG-33	cobalamin-binding	KUO39631.1	Archaea
Hadesarchaea_archaeon_YNP_45	hypothetical	KUO39615.1	Archaea
Hadesarchaea_archaeon_YNP_45	hypothetical	KUO40501.1	Archaea
Hadesarchaea_archaeon_YNP_45	hypothetical	KUO40862.1	Archaea
Hadesarchaea_archaeon_YNP_45	hypothetical	KUO41281.1	Archaea
Hadesarchaea_archaeon_YNP_45	hypothetical	KUO42471.1	Archaea
Hadesarchaea_archaeon_YNP_N21	hypothetical	KUO42253.1	Archaea
Hadesarchaea_archaeon_YNP_N21	hypothetical	KUO42254.1	Archaea
Hadesarchaea_archaeon_YNP_N21	hypothetical	KUO42453.1	Archaea
Hadesarchaea_archaeon_YNP_N21	hypothetical	KUO42842.1	Archaea
Hadesarchaea_archaeon_YNP_N21	hypothetical	KUO43215.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
haloarchaeon_HTSR1	hypothetical	WP_070365541.1	Archaea
haloarchaeon_HTSR1	hypothetical	WP_070365577.1	Archaea
Ignavibacteria_bacterium_RIFCSPLOWO2_12_FULL_56_21	methyltransferase	OGU69916.1	Bacteria
Ignisphaera_aggregans_DSM_17230	cobalamin	ADM27217.1	Archaea
Lokiarchaeum_sp_GC14_75	Dimethylamine	KKK43738.1	Archaea
Lokiarchaeum_sp_GC14_75	Dimethylamine	KKK43983.1	Archaea
Lokiarchaeum_sp_GC14_75	Dimethylamine	KKK44303.1	Archaea
Lokiarchaeum_sp_GC14_75	Dimethylamine	KKK44366.1	Archaea
Lokiarchaeum_sp_GC14_75	putative	KKK45155.1	Archaea
Lokiarchaeum_sp_GC14_75	Trimethylamine	KKK41955.1	Archaea
Lokiarchaeum_sp_GC14_75	Trimethylamine	KKK42684.1	Archaea
Marine_Group_I_thaumarchaeote_SCGC_AAA799-D11	methionine	WP_048089629.1	Archaea
Marine_Group_I_thaumarchaeote_SCGC_AAA799-E16	Methionine	KER05431.1	Archaea
Marine_Group_I_thaumarchaeote_SCGC_AAA799-P11	methionine	WP_048070658.1	Archaea
Marine_Group_I_thaumarchaeote_SCGC_RSA3	MULTISPECIES_	WP_048079445.1	Archaea
Methanobacterium_congolense	Methanol-corrinoid	SCG85713.1	Archaea
Methanobacterium_lacus	cobalamin-binding	WP_013644620.1	Archaea
Methanobacterium_paludis	cobalamin-binding	WP_013825604.1	Archaea
Methanobacterium_sp_A39	cobalamin-binding	WP_069585577.1	Archaea
Methanobacterium	MULTISPECIES_	WP_048081959.1	Archaea
Methanobrevibacter_arboriphilus	cobalamin-binding	WP_042704212.1	Archaea
Methanobrevibacter_smithii_CAG_186	cobalamin-binding	WP_004036708.1	Archaea
Methanobrevibacter_smithii	cobalamin-binding	WP_019262346.1	Archaea
Methanobrevibacter_smithii	cobalamin-binding	WP_019264392.1	Archaea
Methanobrevibacter_sp_A54	MULTISPECIES_	WP_011953942.1	Archaea
Methanobrevibacter_wolinii	cobalamin-binding	WP_042707833.1	Archaea
Methanocella_arvoryzae_MRE50	methanol-cobalamin	WP_012035439.1	Archaea
Methanococcoides_burtonii_DSM_6242	dimethylamine	ABE52282.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanococcoides_burtonii_DSM_6242	dimethylamine	WP_011499430.1	Archaea
Methanococcoides_burtonii_DSM_6242	dimethylamine	WP_011500284.1	Archaea
Methanococcoides_burtonii_DSM_6242	dimethylamine	WP_011500301.1	Archaea
Methanococcoides_burtonii_DSM_6242	methanol-5-hydroxybenzimidazolylcobamide	WP_011498920.1	Archaea
Methanococcoides_burtonii_DSM_6242	methanol-5-hydroxybenzimidazolylcobamide	WP_011498922.1	Archaea
Methanococcoides_burtonii_DSM_6242	monomethylamine	WP_011498946.1	Archaea
Methanococcoides_burtonii_DSM_6242	monomethylamine	WP_011498951.1	Archaea
Methanococcoides_burtonii	dimethylamine	WP_048063580.1	Archaea
Methanococcoides_methylutens_MM1	corrinoid	WP_048205741.1	Archaea
Methanococcoides_methylutens_MM1	Dimethylamine	AKB84490.1	Archaea
Methanococcoides_methylutens_MM1	Dimethylamine	AKB85909.1	Archaea
Methanococcoides_methylutens_MM1	dimethylamine	WP_048205942.1	Archaea
Methanococcoides_methylutens_MM1	hypothetical	WP_048204904.1	Archaea
Methanococcoides_methylutens_MM1	methanol-5-hydroxybenzimidazolylcobamide	WP_048205138.1	Archaea
Methanococcoides_methylutens_MM1	methanol-5-hydroxybenzimidazolylcobamide	WP_048205140.1	Archaea
Methanococcoides_methylutens_MM1	methanol-5-hydroxybenzimidazolylcobamide	WP_048205142.1	Archaea
Methanococcoides_methylutens_MM1	methyltransferase	WP_048206113.1	Archaea
Methanococcoides_methylutens_MM1	monomethylamine	WP_048205165.1	Archaea
Methanococcoides_methylutens_MM1	monomethylamine	WP_048205175.1	Archaea
Methanococcoides_methylutens	dimethylamine	WP_048193756.1	Archaea
Methanococcoides_methylutens	dimethylamine	WP_048195446.1	Archaea
Methanococcoides_methylutens	dimethylamine	WP_048195452.1	Archaea
Methanococcoides_methylutens	dimethylamine	WP_048204688.1	Archaea
Methanococcoides_methylutens	dimethylamine	WP_048205939.1	Archaea
Methanococcoides_methylutens	hypothetical	WP_048194676.1	Archaea
Methanococcoides_methylutens	methanol-5-hydroxybenzimidazolylcobamide	WP_048193144.1	Archaea
Methanococcoides_methylutens	methanol-5-hydroxybenzimidazolylcobamide	WP_048193146.1	Archaea
Methanococcoides_methylutens	methyltransferase	WP_048195756.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanococcoides_methylutens	monomethylamine	WP_048192926.1	Archaea
Methanococcoides_methylutens	monomethylamine	WP_048194174.1	Archaea
Methanococcoides_vulcani	dimethylamine	SES64320.1	Archaea
Methanococcoides_vulcani	dimethylamine	SES80737.1	Archaea
Methanococcoides_vulcani	methanol	SES92369.1	Archaea
Methanococcoides_vulcani	methanol	SET01972.1	Archaea
Methanococcoides_vulcani	methanol	SET02010.1	Archaea
Methanococcoides_vulcani	monomethylamine	SET12072.1	Archaea
Methanococcoides_vulcani	trimethylamine	SES64274.1	Archaea
Methanococcoides_vulcani	trimethylamine	SES92425.1	Archaea
Methanococcus_aeolicus_Nankai-3	cobalamin-binding	WP_011973314.1	Archaea
Methanococcus_maripaludis_C5	cobalamin-binding	WP_011868603.1	Archaea
Methanococcus_maripaludis_C6	cobalamin-binding	WP_012194551.1	Archaea
Methanococcus_maripaludis_C7	cobalamin-binding	WP_011976482.1	Archaea
Methanococcus_maripaludis_S2	Coenzyme	CAF30707.1	Archaea
Methanococcus_maripaludis_X1	cobalamin-binding	WP_011170773.1	Archaea
Methanococcus_maripaludis	cobalamin-binding	WP_048064099.1	Archaea
Methanococcus_voltae_A3	cobalamin-binding	WP_013180957.1	Archaea
Methanococcus_voltae_PS	putative	AAQ55470.1	Archaea
Methanococcus_voltae	putative	CAE02688.1	Archaea
Methanoculleus_sp_CAG_1088	dimethylamine	WP_015505011.1	Archaea
Methanoculleus_sp_CAG_1088	methanol	WP_015504829.1	Archaea
Methanoculleus_sp_CAG_1088	methionine	CDF30150.1	Archaea
Methanoculleus_sp_CAG_1088	methyltransferase	CDF30136.1	Archaea
Methanoculleus_sp_CAG_1088	methyltransferase	CDF30939.1	Archaea
Methanoculleus_sp_CAG_1088	methyltransferase	CDF31086.1	Archaea
Methanoculleus_sp_CAG_1088	trimethylamine	CDF30944.1	Archaea
methanogenic_archaeon_ISO4-H5	dimethylamine	WP_066074949.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
methanogenic_archaeon_ISO4-H5	dimethylamine	WP_066077279.1	Archaea
methanogenic_archaeon_ISO4-H5	dimethylamine	WP_066077287.1	Archaea
methanogenic_archaeon_ISO4-H5	hypothetical	WP_066073323.1	Archaea
methanogenic_archaeon_ISO4-H5	hypothetical	WP_066077282.1	Archaea
methanogenic_archaeon_ISO4-H5	methanol-corrinoid	WP_066075119.1	Archaea
methanogenic_archaeon_ISO4-H5	methanol	AMH94673.1	Archaea
methanogenic_archaeon_ISO4-H5	methionine	WP_066075223.1	Archaea
methanogenic_archaeon_ISO4-H5	methylthiol-CoM	WP_066075225.1	Archaea
methanogenic_archaeon_ISO4-H5	methyltransferase	AMH95045.1	Archaea
methanogenic_archaeon_mixed_culture_ISO4-G1	dimethylamine	AMK13783.1	Archaea
methanogenic_archaeon_mixed_culture_ISO4-G1	dimethylamine	AMK14261.1	Archaea
methanogenic_archaeon_mixed_culture_ISO4-G1	methanol	AMK14086.1	Archaea
methanogenic_archaeon_mixed_culture_ISO4-G1	monomethylamine	AMK14262.1	Archaea
methanogenic_archaeon_mixed_culture_ISO4-G1	trimethylamine	AMK14256.1	Archaea
Methanogenium_cariaci	cobalamin-binding	WP_062396147.1	Archaea
Methanohalobium_evestigatum_Z-7303	cobalamin	WP_013194479.1	Archaea
Methanohalobium_evestigatum_Z-7303	dimethylamine	WP_013194859.1	Archaea
Methanohalobium_evestigatum_Z-7303	dimethylamine	WP_013195341.1	Archaea
Methanohalobium_evestigatum_Z-7303	dimethylamine	WP_013195345.1	Archaea
Methanohalobium_evestigatum_Z-7303	methanol-5-hydroxybenzimidazolylcobamide	WP_013194407.1	Archaea
Methanohalobium_evestigatum_Z-7303	methyltransferase	WP_013195113.1	Archaea
Methanohalobium_evestigatum_Z-7303	monomethylamine	WP_013194522.1	Archaea
Methanohalobium_evestigatum_Z-7303	monomethylamine	WP_013194774.1	Archaea
Methanohalophilus_halophilus	dimethylamine	SDV99323.1	Archaea
Methanohalophilus_halophilus	dimethylamine	SDW40207.1	Archaea
Methanohalophilus_halophilus	dimethylamine	SDW64475.1	Archaea
Methanohalophilus_halophilus	methanol	SDW30294.1	Archaea
Methanohalophilus_halophilus	methanol	SDW30382.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanohalophilus_halophilus	monomethylamine	SDW96043.1	Archaea
Methanohalophilus_halophilus	trimethylamine	SDW64644.1	Archaea
Methanohalophilus_mahii_DSM_5219	cobalamin	WP_013037197.1	Archaea
Methanohalophilus_mahii_DSM_5219	dimethylamine	ADE37171.1	Archaea
Methanohalophilus_mahii_DSM_5219	dimethylamine	WP_013036969.1	Archaea
Methanohalophilus_mahii_DSM_5219	dimethylamine	WP_013038117.1	Archaea
Methanohalophilus_mahii_DSM_5219	methanol-5-hydroxybenzimidazolylcobamide	WP_013037438.1	Archaea
Methanohalophilus_mahii_DSM_5219	methanol-5-hydroxybenzimidazolylcobamide	WP_013037440.1	Archaea
Methanohalophilus_mahii_DSM_5219	methyltransferase	WP_013037920.1	Archaea
Methanohalophilus_mahii_DSM_5219	monomethylamine	WP_013037586.1	Archaea
Methanohalophilus_mahii	dimethylamine	WP_048902185.1	Archaea
Methanohalophilus_portucalensis_FDF-1	corrinoid	ABQ44361.1	Archaea
Methanohalophilus_sp_2-GBenrich	methanol-5-hydroxybenzimidazolylcobamide	OBZ35203.1	Archaea
Methanohalophilus_sp_2-GBenrich	methanol	ODV49342.1	Archaea
Methanohalophilus_sp_2-GBenrich	trimethylamine	KXS43947.1	Archaea
Methanohalophilus_sp_DAL1	dimethylamine	OBZ34966.1	Archaea
Methanohalophilus_sp_DAL1	dimethylamine	OBZ35608.1	Archaea
Methanohalophilus_sp_DAL1	methanol-5-hydroxybenzimidazolylcobamide	OBZ35201.1	Archaea
Methanohalophilus_sp_T328-1	methanol	KXS39624.1	Archaea
Methanohalophilus_sp_T328-1	methanol	KXS41359.1	Archaea
Methanolacinia_paynteri	cobalamin-binding	WP_052418638.1	Archaea
Methanolacinia_petrolearia_DSM_11571	cobalamin-binding	WP_013330474.1	Archaea
Methanolobus_profundi	dimethylamine	SFM70936.1	Archaea
Methanolobus_profundi	dimethylamine	SFM71364.1	Archaea
Methanolobus_profundi	methanol	SFM28187.1	Archaea
Methanolobus_profundi	methanol	SFM28225.1	Archaea
Methanolobus_profundi	methanol	SFM60632.1	Archaea
Methanolobus_profundi	methylmalonyl-CoA	SFM31281.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanolobus_profundi	monomethylamine	SFM70883.1	Archaea
Methanolobus_profundi	monomethylamine	SFM71262.1	Archaea
Methanolobus_profundi	trimethylamine	SFM71036.1	Archaea
Methanolobus_psychrophilus_R15	dimethylamine	WP_015053602.1	Archaea
Methanolobus_psychrophilus_R15	dimethylamine	WP_015053607.1	Archaea
Methanolobus_psychrophilus_R15	dimethylamine	WP_015053623.1	Archaea
Methanolobus_psychrophilus_R15	methanol-5-hydroxybenzimidazolylcobamide	WP_015052840.1	Archaea
Methanolobus_psychrophilus_R15	methanol-5-hydroxybenzimidazolylcobamide	WP_015054929.1	Archaea
Methanolobus_psychrophilus_R15	methanol	AFV25234.1	Archaea
Methanolobus_psychrophilus_R15	monomethylamine	WP_015053599.1	Archaea
Methanolobus_psychrophilus_R15	monomethylamine	WP_015053634.1	Archaea
Methanolobus_psychrophilus	methanol-5-hydroxybenzimidazolylcobamide	WP_048147693.1	Archaea
Methanolobus_sp_T82-4	Dimethylamine	KXS40258.1	Archaea
Methanolobus_sp_T82-4	methanol	KXS42386.1	Archaea
Methanolobus_sp_T82-4	methanol	KXS42388.1	Archaea
Methanolobus_sp_T82-4	methanol	KXS44159.1	Archaea
Methanolobus_sp_T82-4	methyltransferase	KXS42690.1	Archaea
Methanolobus_sp_T82-4	monomethylamine	KXS40255.1	Archaea
Methanolobus_sp_T82-4	monomethylamine	KXS40347.1	Archaea
Methanolobus_sp_T82-4	putative	KXS41029.1	Archaea
Methanolobus_sp_T82-4	trimethylamine	KXS41813.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023844869.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023845219.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023845231.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023845237.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023845247.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023845251.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023845253.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023846065.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023846488.1	Archaea
Methanolobus_tindarius_DSM_2278	methyltransferase	WP_023846490.1	Archaea
Methanolobus_tindarius_DSM_2278	putative	ETA67038.1	Archaea
Methanolobus_vulcani	B12	SDG36752.1	Archaea
Methanolobus_vulcani	dimethylamine	SDG39443.1	Archaea
Methanolobus_vulcani	methanol	SDF30409.1	Archaea
Methanolobus_vulcani	methanol	SDG14619.1	Archaea
Methanolobus_vulcani	methanol	SDG14676.1	Archaea
Methanolobus_vulcani	methylmalonyl-CoA	SDG04186.1	Archaea
Methanolobus_vulcani	monomethylamine	SDF92866.1	Archaea
Methanolobus_vulcani	monomethylamine	SDG08857.1	Archaea
Methanolobus_vulcani	monomethylamine	SDG39511.1	Archaea
Methanolobus_vulcani	trimethylamine	SDF80410.1	Archaea
Methanolobus_vulcani	trimethylamine	SDG39347.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	dimethylamine	KQM11577.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	hypothetical	KQM11572.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	methanol-corrinoid	KQM12197.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	methanol-corrinoid	KQM12673.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	methanol-corrinoid	KQM12747.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	methionine	KQM11929.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	methionine	KQM11939.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M1	methylthiol-CoM	KQM12402.1	Archaea
Methanomassiliicoccales_archaeon_RumEn_M2	hypothetical	KQM10800.1	Archaea
Methanomassiliicoccus_luminyensis	dimethylamine	WP_026068678.1	Archaea
Methanomassiliicoccus_luminyensis	dimethylamine	WP_026068679.1	Archaea
Methanomassiliicoccus_luminyensis	dimethylamine	WP_026069100.1	Archaea
Methanomassiliicoccus_luminyensis	dimethylamine	WP_026069102.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanomassiliicoccus_luminyensis	dimethylamine	WP_049796371.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019176359.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019176763.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019176991.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019177442.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019177725.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019177907.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019178518.1	Archaea
Methanomassiliicoccus_luminyensis	hypothetical	WP_019178649.1	Archaea
Methanomassiliicoccus_luminyensis	methanol-corrinoid	WP_026069014.1	Archaea
Methanomassiliicoccus_luminyensis	methylthiol-CoM	WP_026068850.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	cobalamin-binding	WP_015323796.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	cobalamin-binding	WP_015324657.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	cobalamin-binding	WP_015324938.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	methyltransferase	WP_015323483.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	methyltransferase	WP_015323497.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	methyltransferase	WP_015323501.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	methyltransferase	WP_015324190.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	methyltransferase	WP_015324960.1	Archaea
Methanomethylovorans_hollandica_DSM_15978	methyltransferase	WP_015324961.1	Archaea
Methanomicrobiales_archaeon_53_19	5-methyltetrahydrofolate-homocysteinemethyltrans	KUL00639.1	Archaea
Methanoplanus_limicola_DSM_2279	cobalamin-binding	WP_004077249.1	Archaea
Methanoplanus_limicola_DSM_2279	cobalamin-binding	WP_004077272.1	Archaea
Methanoregula_boonei_6A8	cobalamin-binding	WP_012106586.1	Archaea
Methanosalsum_zhilinae_DSM_4017	dimethylamine	WP_013897774.1	Archaea
Methanosalsum_zhilinae_DSM_4017	dimethylamine	WP_013897910.1	Archaea
Methanosalsum_zhilinae_DSM_4017	methanol-5-hydroxybenzimidazolylcobamide	WP_013899222.1	Archaea
Methanosalsum_zhilinae_DSM_4017	methanol-5-hydroxybenzimidazolylcobamide	WP_013899224.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosalsum_zhilinae_DSM_4017	methyltransferase	AEH61362.1	Archaea
Methanosalsum_zhilinae_DSM_4017	methyltransferase	WP_013897673.1	Archaea
Methanosalsum_zhilinae_DSM_4017	methyltransferase	WP_013899066.1	Archaea
Methanosalsum_zhilinae_DSM_4017	monomethylamine	WP_013897935.1	Archaea
Methanosalsum_zhilinae_DSM_4017	monomethylamine	WP_013898801.1	Archaea
Methanosarcina_acetivorans_C2A	corrinoid	AAM07706.1	Archaea
Methanosarcina_acetivorans_C2A	corrinoid	WP_011024052.1	Archaea
Methanosarcina_acetivorans_C2A	dimethylamine	AAM04367.1	Archaea
Methanosarcina_acetivorans_C2A	dimethylamine	WP_011020576.1	Archaea
Methanosarcina_acetivorans_C2A	dimethylamine	WP_011020578.1	Archaea
Methanosarcina_acetivorans_C2A	dimethylamine	WP_011020969.1	Archaea
Methanosarcina_acetivorans_C2A	dimethylamine	WP_011022395.1	Archaea
Methanosarcina_acetivorans_C2A	hypothetical	AAM04298.1	Archaea
Methanosarcina_acetivorans_C2A	methanol-5-hydroxybenzimidazolylcobamide	WP_011020507.1	Archaea
Methanosarcina_acetivorans_C2A	methanol-5-hydroxybenzimidazolylcobamide	WP_011021627.1	Archaea
Methanosarcina_acetivorans_C2A	methanol-5-hydroxybenzimidazolylcobamide	WP_011024270.1	Archaea
Methanosarcina_acetivorans_C2A	methyltransferase	WP_011024263.1	Archaea
Methanosarcina_acetivorans_C2A	methyltransferase	WP_011024431.1	Archaea
Methanosarcina_acetivorans_C2A	monomethylamine	WP_011020203.1	Archaea
Methanosarcina_acetivorans_C2A	monomethylamine	WP_011022912.1	Archaea
Methanosarcina_acetivorans	corrinoid	WP_048066589.1	Archaea
Methanosarcina_acetivorans	dimethylamine	WP_048065003.1	Archaea
Methanosarcina_acetivorans	methyltransferase	WP_048064984.1	Archaea
Methanosarcina_barkeri_227	Dimethylamine	AKB53925.1	Archaea
Methanosarcina_barkeri_227	methylthiol-CoM	WP_048117949.1	Archaea
Methanosarcina_barkeri_227	Monomethylamine	AKB58746.1	Archaea
Methanosarcina_barkeri_3	corrinoid	WP_048107480.1	Archaea
Methanosarcina_barkeri_3	Dimethylamine	AKB82528.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_barkeri_3	dimethylamine	WP_048106279.1	Archaea
Methanosarcina_barkeri_3	dimethylamine	WP_048106458.1	Archaea
Methanosarcina_barkeri_3	dimethylamine	WP_048108036.1	Archaea
Methanosarcina_barkeri_3	methanol-5-hydroxybenzimidazolylcobamide	WP_048108215.1	Archaea
Methanosarcina_barkeri_3	methanol-5-hydroxybenzimidazolylcobamide	WP_048109505.1	Archaea
Methanosarcina_barkeri_3	Methanol	AKB82957.1	Archaea
Methanosarcina_barkeri_3	methylthiol-CoM	WP_048107235.1	Archaea
Methanosarcina_barkeri_3	monomethylamine	WP_048106316.1	Archaea
Methanosarcina_barkeri_3	monomethylamine	WP_048108466.1	Archaea
Methanosarcina_barkeri_3	monomethylamine	WP_048108472.1	Archaea
Methanosarcina_barkeri_CM1	corrinoid	WP_048176401.1	Archaea
Methanosarcina_barkeri_CM1	Dimethylamine	AKB53460.1	Archaea
Methanosarcina_barkeri_CM1	dimethylamine	WP_048177707.1	Archaea
Methanosarcina_barkeri_CM1	methyltransferase	AKJ39555.1	Archaea
Methanosarcina_barkeri_str_Fusaro	methanol-5-hydroxybenzimidazolylcobamide	WP_011306079.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	corrinoid	WP_011307117.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	dimethylamine	AAZ70456.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	dimethylamine	WP_011306503.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	dimethylamine	WP_011308511.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	dimethylamine	WP_048136559.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	methanol-5-hydroxybenzimidazolylcobamide	WP_011305768.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	methanol	AAZ72501.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	Methanol	AKB52912.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	methylthiol-coenzyme	WP_011308491.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	monomethylamine	WP_011305868.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	monomethylamine	WP_011305870.1	Archaea
Methanosarcina_barkeri_str_Wiesmoor	monomethylamine	WP_011308537.1	Archaea
Methanosarcina_barkeri	dimethylamine	WP_048102697.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_barkeri	dimethylamine	WP_048108033.1	Archaea
Methanosarcina_barkeri	methanol-5-hydroxybenzimidazolylcobamide	WP_048103078.1	Archaea
Methanosarcina_barkeri	methanol-5-hydroxybenzimidazolylcobamide	WP_048108608.1	Archaea
Methanosarcina_barkeri	monomethylamine	WP_048123347.1	Archaea
Methanosarcina_barkeri	monomethylamine	WP_048177688.1	Archaea
Methanosarcina_barkeri	RecName_	Q9P9L5.1	Archaea
Methanosarcina_flavescens	corrinoid	WP_054297895.1	Archaea
Methanosarcina_flavescens	dimethylamine	WP_054297638.1	Archaea
Methanosarcina_flavescens	dimethylamine	WP_054298262.1	Archaea
Methanosarcina_flavescens	methanol-5-hydroxybenzimidazolylcobamide	WP_054297871.1	Archaea
Methanosarcina_flavescens	methanol-5-hydroxybenzimidazolylcobamide	WP_054299618.1	Archaea
Methanosarcina_flavescens	methanol-5-hydroxybenzimidazolylcobamide	WP_054299961.1	Archaea
Methanosarcina_flavescens	monomethylamine	WP_054298912.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	Dimethylamine	AKB78616.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	Dimethylamine	AKB80173.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	dimethylamine	WP_048138814.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	dimethylamine	WP_048142237.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	hypothetical	WP_052730663.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	methanol-5-hydroxybenzimidazolylcobamide	WP_048136827.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	methanol-5-hydroxybenzimidazolylcobamide	WP_048138492.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	methanol-5-hydroxybenzimidazolylcobamide	WP_048142319.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	methyltransferase	WP_048136830.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	methyltransferase	WP_048141746.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	Monomethylamine	AKB78301.1	Archaea
Methanosarcina_horonobensis_HB-1__JCM_15518	monomethylamine	WP_048142712.1	Archaea
Methanosarcina_horonobensis	dimethylamine	WP_048142243.1	Archaea
Methanosarcina_horonobensis	dimethylamine	WP_048143406.1	Archaea
Methanosarcina_horonobensis	monomethylamine	WP_048143342.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_lacustris_Z-7289	corrinoid	WP_048127909.1	Archaea
Methanosarcina_lacustris_Z-7289	Dimethylamine	AKB73547.1	Archaea
Methanosarcina_lacustris_Z-7289	dimethylamine	WP_048124402.1	Archaea
Methanosarcina_lacustris_Z-7289	dimethylamine	WP_048127482.1	Archaea
Methanosarcina_lacustris_Z-7289	dimethylamine	WP_048127492.1	Archaea
Methanosarcina_lacustris_Z-7289	dimethylamine	WP_048129223.1	Archaea
Methanosarcina_lacustris_Z-7289	methanol-5-hydroxybenzimidazolylcobamide	WP_048124347.1	Archaea
Methanosarcina_lacustris_Z-7289	methanol-5-hydroxybenzimidazolylcobamide	WP_048125662.1	Archaea
Methanosarcina_lacustris_Z-7289	methanol-5-hydroxybenzimidazolylcobamide	WP_048128320.1	Archaea
Methanosarcina_lacustris_Z-7289	methylthiol-CoM	WP_048124885.1	Archaea
Methanosarcina_lacustris_Z-7289	methyltransferase	WP_048127715.1	Archaea
Methanosarcina_lacustris_Z-7289	methyltransferase	WP_048127724.1	Archaea
Methanosarcina_lacustris_Z-7289	methyltransferase	WP_048128332.1	Archaea
Methanosarcina_lacustris_Z-7289	monomethylamine	WP_048128728.1	Archaea
Methanosarcina_lacustris_Z-7289	monomethylamine	WP_048128733.1	Archaea
Methanosarcina_lacustris	dimethylamine	WP_048124399.1	Archaea
Methanosarcina_mazei_C16	monomethylamine	WP_048042243.1	Archaea
Methanosarcina_mazei_C16	RecName_	P58981.1	Archaea
Methanosarcina_mazei_Go1	dimethylamine	AAM31748.1	Archaea
Methanosarcina_mazei_Go1	dimethylamine	AAM32657.1	Archaea
Methanosarcina_mazei_Go1	Methanol	AAM29870.1	Archaea
Methanosarcina_mazei_Go1	Methanol	AAM31344.1	Archaea
Methanosarcina_mazei_Go1	Monomethylamine	AAM31134.1	Archaea
Methanosarcina_mazei_Go1	trimethylamine	AAM30751.1	Archaea
Methanosarcina_mazei_LYC	Dimethylamine	AKB41894.1	Archaea
Methanosarcina_mazei_LYC	Methanol	AKB39227.1	Archaea
Methanosarcina_mazei_LYC	monomethylamine	WP_011035223.1	Archaea
Methanosarcina_mazei_LYC	MULTISPECIES_	WP_048041112.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_mazei_S-6	Corrinoid	AKB63440.1	Archaea
Methanosarcina_mazei_SarPi	methylthiol-CoM	WP_048043670.1	Archaea
Methanosarcina_mazei_SarPi	monomethylamine	WP_048043122.1	Archaea
Methanosarcina_mazei_Tuc01	Dimethylamine	AGF97112.1	Archaea
Methanosarcina_mazei_Tuc01	Dimethylamine	AGF97446.1	Archaea
Methanosarcina_mazei_Tuc01	Methanol	AGF96500.1	Archaea
Methanosarcina_mazei_Tuc01	Monomethylamine	AGF98689.1	Archaea
Methanosarcina_mazei_Tuc01	Trimethylamine	AGF97114.1	Archaea
Methanosarcina_mazei_Tuc01	Trimethylamine	AGF97445.1	Archaea
Methanosarcina_mazei	corrinoid	WP_048036675.1	Archaea
Methanosarcina_mazei	corrinoid	WP_048046902.1	Archaea
Methanosarcina_mazei	dimethylamine	KKH ₂ 6279.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_011033632.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048036696.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048037232.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048039209.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048039218.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048039288.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048040699.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048041223.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048041924.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048042651.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048043719.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048045614.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048047172.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048048845.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048049098.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048049720.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_mazei	dimethylamine	WP_048049733.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048049918.1	Archaea
Methanosarcina_mazei	dimethylamine	WP_048049923.1	Archaea
Methanosarcina_mazei	methanol-5-hydroxybenzimidazolylcobamide	WP_048038194.1	Archaea
Methanosarcina_mazei	methanol-5-hydroxybenzimidazolylcobamide	WP_048046707.1	Archaea
Methanosarcina_mazei	methanol	WP_015410915.1	Archaea
Methanosarcina_mazei	methylthiol-coenzyme	WP_011034352.1	Archaea
Methanosarcina_mazei	monomethylamine	KKG05940.1	Archaea
Methanosarcina_mazei	monomethylamine	KKG08075.1	Archaea
Methanosarcina_mazei	monomethylamine	KKG35840.1	Archaea
Methanosarcina_mazei	monomethylamine	KKG71955.1	Archaea
Methanosarcina_mazei	monomethylamine	KKG84684.1	Archaea
Methanosarcina_mazei	monomethylamine	KKG90767.1	Archaea
Methanosarcina_mazei	monomethylamine	KKH ₂ 2756.1	Archaea
Methanosarcina_mazei	monomethylamine	KKH ₂ 8501.1	Archaea
Methanosarcina_mazei	monomethylamine	KKH30479.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048041204.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048041216.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048043677.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048045046.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048045194.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048046887.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048047888.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048047962.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048049148.1	Archaea
Methanosarcina_mazei	monomethylamine	WP_048049248.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_011033022.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_011033979.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_mazei	MULTISPECIES_	WP_015411537.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_015411985.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_015412949.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_048036790.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_048036796.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_048037606.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_048039968.1	Archaea
Methanosarcina_mazei	MULTISPECIES_	WP_048046403.1	Archaea
Methanosarcina_siciliae_C2J	Corrinoid	AKB38646.1	Archaea
Methanosarcina_siciliae_C2J	corrinoid	WP_052727389.1	Archaea
Methanosarcina_siciliae_C2J	Dimethylamine	AKB27170.1	Archaea
Methanosarcina_siciliae_C2J	Dimethylamine	AKB36153.1	Archaea
Methanosarcina_siciliae_C2J	dimethylamine	WP_048169686.1	Archaea
Methanosarcina_siciliae_C2J	dimethylamine	WP_048172322.1	Archaea
Methanosarcina_siciliae_C2J	methanol-5-hydroxybenzimidazolylcobamide	WP_048169586.1	Archaea
Methanosarcina_siciliae_C2J	methanol-5-hydroxybenzimidazolylcobamide	WP_048174076.1	Archaea
Methanosarcina_siciliae_C2J	methanol-5-hydroxybenzimidazolylcobamide	WP_048181634.1	Archaea
Methanosarcina_siciliae_C2J	methyltransferase	WP_048180102.1	Archaea
Methanosarcina_siciliae_C2J	methyltransferase	WP_048185089.1	Archaea
Methanosarcina_siciliae_C2J	methyltransferase	WP_048185291.1	Archaea
Methanosarcina_siciliae_C2J	monomethylamine	WP_048169156.1	Archaea
Methanosarcina_siciliae_C2J	monomethylamine	WP_048171675.1	Archaea
Methanosarcina_siciliae_C2J	monomethylamine	WP_048178969.1	Archaea
Methanosarcina_siciliae_HI350	Corrinoid	AKB34282.1	Archaea
Methanosarcina_siciliae_HI350	corrinoid	WP_052721721.1	Archaea
Methanosarcina_siciliae_HI350	dimethylamine	WP_048171401.1	Archaea
Methanosarcina_siciliae_HI350	methanol-5-hydroxybenzimidazolylcobamide	WP_048171031.1	Archaea
Methanosarcina_siciliae_HI350	methyltransferase	WP_048170227.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_siciliae_HI350	methyltransferase	WP_048174070.1	Archaea
Methanosarcina_siciliae_HI350	methyltransferase	WP_048174258.1	Archaea
Methanosarcina_siciliae_HI350	monomethylamine	WP_048169165.1	Archaea
Methanosarcina_siciliae	corrinoid	WP_048175113.1	Archaea
Methanosarcina_siciliae	corrinoid	WP_048186226.1	Archaea
Methanosarcina_siciliae	dimethylamine	WP_048169681.1	Archaea
Methanosarcina_soligelidi	methanol-5-hydroxybenzimidazolylcobamide	WP_048051775.1	Archaea
Methanosarcina_sp_1_H_A_2_2	corrinoid	WP_048161040.1	Archaea
Methanosarcina_sp_1_H_A_2_2	dimethylamine	WP_048162583.1	Archaea
Methanosarcina_sp_1_H_A_2_2	dimethylamine	WP_048162859.1	Archaea
Methanosarcina_sp_1_H_A_2_2	dimethylamine	WP_048162860.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methanol-5-hydroxybenzimidazolylcobamide	WP_048161310.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methylthiol-CoM	WP_048161506.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methyltransferase	WP_048161132.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methyltransferase	WP_048161137.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methyltransferase	WP_048161314.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methyltransferase	WP_048162492.1	Archaea
Methanosarcina_sp_1_H_A_2_2	methyltransferase	WP_048162494.1	Archaea
Methanosarcina_sp_1_H_A_2_2	monomethylamine	WP_048160756.1	Archaea
Methanosarcina_sp_1_H_A_2_2	MULTISPECIES_	WP_048137022.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	corrinoid	WP_048135972.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	dimethylamine	WP_048135050.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	dimethylamine	WP_048135208.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	methylthiol-CoM	WP_048132558.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	methyltransferase	WP_048131792.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	methyltransferase	WP_048131802.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	methyltransferase	WP_048133540.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	methyltransferase	WP_048136429.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_sp_1_H_T_1A_1	methyltransferase	WP_048136495.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	monomethylamine	WP_048133385.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048132529.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048132713.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048133189.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048133545.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048134375.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048134381.1	Archaea
Methanosarcina_sp_1_H_T_1A_1	MULTISPECIES_	WP_048134867.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	corrinoid	WP_048169974.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	dimethylamine	WP_048171312.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	dimethylamine	WP_048171528.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	dimethylamine	WP_048172600.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	methanol-5-hydroxybenzimidazolylcobamide	WP_048169521.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	methylthiol-CoM	WP_048169094.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	methyltransferase	WP_048171866.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	methyltransferase	WP_048173052.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	methyltransferase	WP_048173060.1	Archaea
Methanosarcina_sp_2_H_A_1B_4	monomethylamine	WP_048169770.1	Archaea
Methanosarcina_sp_2_H_T_1A_15	dimethylamine	WP_048144146.1	Archaea
Methanosarcina_sp_2_H_T_1A_15	hypothetical	KKG28724.1	Archaea
Methanosarcina_sp_2_H_T_1A_15	monomethylamine	KKG14709.1	Archaea
Methanosarcina_sp_2_H_T_1A_15	MULTISPECIES_	WP_048138760.1	Archaea
Methanosarcina_sp_2_H_T_1A_15	MULTISPECIES_	WP_048142974.1	Archaea
Methanosarcina_sp_2_H_T_1A_3	monomethylamine	KKG19060.1	Archaea
Methanosarcina_sp_2_H_T_1A_3	monomethylamine	WP_048141722.1	Archaea
Methanosarcina_sp_2_H_T_1A_6	MULTISPECIES_	WP_048137445.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	methanol-5-hydroxybenzimidazolylcobamide	KKG15808.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_sp_2_H_T_1A_8	monomethylamine	KKG20171.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048137826.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048137836.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048137869.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048140300.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048140304.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048140795.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048142837.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048142961.1	Archaea
Methanosarcina_sp_2_H_T_1A_8	MULTISPECIES_	WP_048144042.1	Archaea
Methanosarcina_sp_795	dimethylamine	WP_048166440.1	Archaea
Methanosarcina_sp_795	dimethylamine	WP_048167188.1	Archaea
Methanosarcina_sp_795	dimethylamine	WP_048167193.1	Archaea
Methanosarcina_sp_795	dimethylamine	WP_048167810.1	Archaea
Methanosarcina_sp_795	methanol-5-hydroxybenzimidazolylcobamide	WP_048166109.1	Archaea
Methanosarcina_sp_795	methanol-5-hydroxybenzimidazolylcobamide	WP_048167521.1	Archaea
Methanosarcina_sp_795	methanol-5-hydroxybenzimidazolylcobamide	WP_048167874.1	Archaea
Methanosarcina_sp_795	monomethylamine	WP_048167688.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048117364.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048117958.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048119022.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048120097.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048120103.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048120367.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048120781.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048122309.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048123111.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048154921.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_sp_A14	MULTISPECIES_	WP_048155349.1	Archaea
Methanosarcina_sp_A14	MULTISPECIES_	WP_048156643.1	Archaea
Methanosarcina_sp_Ant1	dimethylamine	OEU42523.1	Archaea
Methanosarcina_sp_Ant1	dimethylamine	OEU42560.1	Archaea
Methanosarcina_sp_Ant1	dimethylamine	OEU42563.1	Archaea
Methanosarcina_sp_Ant1	methanol-5-hydroxybenzimidazolylcobamide	OEU41050.1	Archaea
Methanosarcina_sp_Ant1	methanol-5-hydroxybenzimidazolylcobamide	OEU43385.1	Archaea
Methanosarcina_sp_Ant1	methanol-5-hydroxybenzimidazolylcobamide	OEU43455.1	Archaea
Methanosarcina_sp_Ant1	methylthiol-CoM	OEU43011.1	Archaea
Methanosarcina_sp_Kolksee	Dimethylamine	AKB42746.1	Archaea
Methanosarcina_sp_Kolksee	dimethylamine	WP_048154426.1	Archaea
Methanosarcina_sp_Kolksee	methylthiol-CoM	WP_048155339.1	Archaea
Methanosarcina_sp_Kolksee	monomethylamine	WP_048155792.1	Archaea
Methanosarcina_sp_Kolksee	monomethylamine	WP_048158455.1	Archaea
Methanosarcina_sp_Kolksee	MULTISPECIES_	WP_048116961.1	Archaea
Methanosarcina_sp_Kolksee	MULTISPECIES_	WP_048117314.1	Archaea
Methanosarcina_sp_Kolksee	MULTISPECIES_	WP_048117652.1	Archaea
Methanosarcina_sp_Kolksee	MULTISPECIES_	WP_048119769.1	Archaea
Methanosarcina_sp_Kolksee	MULTISPECIES_	WP_048120666.1	Archaea
Methanosarcina_sp_MTP4	corrinoid	WP_052718253.1	Archaea
Methanosarcina_sp_MTP4	Dimethylamine	AKB26017.1	Archaea
Methanosarcina_sp_MTP4	dimethylamine	WP_048179020.1	Archaea
Methanosarcina_sp_MTP4	dimethylamine	WP_048180568.1	Archaea
Methanosarcina_sp_MTP4	dimethylamine	WP_048180583.1	Archaea
Methanosarcina_sp_MTP4	dimethylamine	WP_048181470.1	Archaea
Methanosarcina_sp_MTP4	dimethylamine	WP_048181479.1	Archaea
Methanosarcina_sp_MTP4	dimethylamine	WP_048183147.1	Archaea
Methanosarcina_sp_MTP4	hypothetical	WP_048177758.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_sp_MTP4	hypothetical	WP_048177762.1	Archaea
Methanosarcina_sp_MTP4	hypothetical	WP_048179054.1	Archaea
Methanosarcina_sp_MTP4	hypothetical	WP_048179060.1	Archaea
Methanosarcina_sp_MTP4	hypothetical	WP_048179071.1	Archaea
Methanosarcina_sp_MTP4	hypothetical	WP_048181820.1	Archaea
Methanosarcina_sp_MTP4	methanol-5-hydroxybenzimidazolylcobamide	WP_048177317.1	Archaea
Methanosarcina_sp_MTP4	methanol-5-hydroxybenzimidazolylcobamide	WP_048178272.1	Archaea
Methanosarcina_sp_MTP4	methanol-5-hydroxybenzimidazolylcobamide	WP_048181508.1	Archaea
Methanosarcina_sp_MTP4	methyltransferase	WP_048177073.1	Archaea
Methanosarcina_sp_MTP4	methyltransferase	WP_048177324.1	Archaea
Methanosarcina_sp_MTP4	methyltransferase	WP_048180685.1	Archaea
Methanosarcina_sp_MTP4	methyltransferase	WP_048180700.1	Archaea
Methanosarcina_sp_MTP4	monomethylamine	WP_048178606.1	Archaea
Methanosarcina_sp_MTP4	monomethylamine	WP_048181993.1	Archaea
Methanosarcina_sp_WH1	corrinoid	WP_048129936.1	Archaea
Methanosarcina_sp_WH1	Dimethylamine	AKB17212.1	Archaea
Methanosarcina_sp_WH1	Dimethylamine	AKB21497.1	Archaea
Methanosarcina_sp_WH1	Methanol	AKB17957.1	Archaea
Methanosarcina_sp_WH1	monomethylamine	WP_048130009.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048125500.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048125544.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048125873.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048125881.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048126966.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048127793.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048129297.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048129310.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048129422.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanosarcina_sp_WH1	MULTISPECIES_	WP_048130008.1	Archaea
Methanosarcina_sp_WH1	MULTISPECIES_	WP_052722545.1	Archaea
Methanosarcina_sp_WWM596	corrinoid	WP_048159063.1	Archaea
Methanosarcina_sp_WWM596	Dimethylamine	AKB18166.1	Archaea
Methanosarcina_thermophila_CHTI-55	Dimethylamine	AKB13820.1	Archaea
Methanosarcina_thermophila_CHTI-55	dimethylamine	WP_048167811.1	Archaea
Methanosarcina_thermophila_CHTI-55	Methanol	AKB11842.1	Archaea
Methanosarcina_thermophila_TM-1	Dimethylamine	AKB13118.1	Archaea
Methanosarcina_vacuolata_Z-761	dimethylamine	WP_048119496.1	Archaea
Methanosarcina_vacuolata_Z-761	methylthiol-CoM	WP_048120317.1	Archaea
Methanosarcina_vacuolata_Z-761	monomethylamine	WP_048120646.1	Archaea
Methanosarcina_vacuolata_Z-761	monomethylamine	WP_048124322.1	Archaea
Methanosarcina	MULTISPECIES_	WP_048117644.1	Archaea
Methanosarcina	MULTISPECIES_	WP_048125538.1	Archaea
Methanosarcina	MULTISPECIES_	WP_048126775.1	Archaea
Methanosarcina	MULTISPECIES_	WP_048140391.1	Archaea
Methanosarcinales_archeaon_56_1174	cobalamin-binding	WP_042685521.1	Archaea
Methanosarcinales_archeaon_56_1174	dimethylamine	WP_042686889.1	Archaea
Methanosarcinales_archeaon_56_1174	hypothetical	WP_042686089.1	Archaea
Methanosarcinales_archeaon_56_1174	Methyltransferase	KUK04366.1	Archaea
Methanosphaera_sp_A6	MULTISPECIES_	WP_011405801.1	Archaea
Methanosphaera_sp_A6	MULTISPECIES_	WP_011405805.1	Archaea
Methanosphaera_sp_WGK6	cobalamin-binding	WP_069592772.1	Archaea
Methanosphaera_sp_WGK6	cobalamin-binding	WP_069592774.1	Archaea
Methanosphaera_sp_WGK6	cobalamin-binding	WP_069593532.1	Archaea
Methanosphaera_stadtmanae_DSM_3091	cobalamin-binding	WP_011405803.1	Archaea
Methanosphaera_stadtmanae_DSM_3091	corrinoid	CAE48302.1	Archaea
Methanosphaerula_palustris_E1-9c	cobalamin-binding	WP_012617957.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Methanothermococcus_thermolithotrophicus	cobalamin-binding	WP_018153949.1	Archaea
Methanotorris_igneus_Kol_5	cobalamin-binding	WP_013798831.1	Archaea
Methermicoccus_shengliensis	dimethylamine	WP_042686907.1	Archaea
Methermicoccus_shengliensis	hypothetical	WP_042684257.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ-80	hypothetical	KON28092.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ-80	hypothetical	KON28826.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ-80	methyltransferase	KON26718.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ1-55	hypothetical	KON26252.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ1-55	hypothetical	KON27240.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ1-55	hypothetical	KON28056.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ1-55	hypothetical	KON30388.1	Archaea
miscellaneous_Crenarchaeota_group_archaeon_SMTZ1-55	hypothetical	KON30703.1	Archaea
miscellaneous_Crenarchaeota_group-15_archaeon_DG-45	hypothetical	KON29503.1	Archaea
miscellaneous_Crenarchaeota_group-15_archaeon_DG-45	hypothetical	KON30756.1	Archaea
Nitriliruptor_alkaliphilus	5-methyltetrahydrofolate-homocysteine	WP_052665966.1	Bacteria
Nitrosopumilus_maritimus_SCM1	Methionine	ABX13163.1	Bacteria
Nitrosopumilus_sp_BACL13_MAG-121220-bin23	methionine	KRO29036.1	Bacteria
Nitrosopumilus_sp_Nsub	methionine	WP_067958663.1	Bacteria
Nitrososphaera_viennensis_EN76	putative	AIC16863.1	Bacteria
Parasporobacterium_paucivorans_DSM_15970	dimethylamine	WP_073994062.1	Bacteria
Pelosinus_sp_UFO1	cobalamin-binding	WP_038671865.1	Bacteria
Peptococcaceae_bacterium_CEB3	dimethylamine	WP_047828021.1	Bacteria
Peptococcaceae_bacterium_SCADC1_2_3	dimethylamine	KFD41411.1	Bacteria
Proteiniborus_sp_DW1	cobalamin-binding	WP_074349034.1	Bacteria
Pseudobacteroides_cellulosolvens_ATCC_35603__DSM_2933	dimethylamine	WP_050753528.1	Bacteria
Pseudobacteroides_cellulosolvens	dimethylamine	WP_036939029.1	Bacteria
Sporomusa_acidovorans	methyltransferase	SDF24799.1	Bacteria
Sporomusa_sp_An4	MULTISPECIES_	WP_021166618.1	Bacteria

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Staphylothermus_hellenicus_DSM_12710	corrinoid	WP_013142718.1	Archaea
Staphylothermus_marinus_F1	corrinoid	WP_011838717.1	Archaea
Syntrophaceticus_schinkii	methyltransferase	WP_044663967.1	Archaea
Syntrophaceticus_schinkii	methyltransferase	WP_044664928.1	Archaea
Syntrophaceticus_schinkii	methyltransferase	WP_044665061.1	Archaea
Syntrophobotulus_glycolicus_DSM_8271	cobalamin-binding	WP_013625686.1	Bacteria
Thaumarchaeota_archaeon_CSP1-1	methionine	KRT61287.1	Archaea
Thaumarchaeota_archaeon_MY2	methionine	WP_042687476.1	Archaea
Thaumarchaeota_archaeon_MY3	Methionine	ALI36678.1	Archaea
Thaumarchaeota_archaeon_N4	methionine	WP_048196937.1	Archaea
Thaumarchaeota_archaeon_RBG_16_49_8	hypothetical	OHE53163.1	Archaea
Thaumarchaeota_archaeon_SCGC_AB-539-E09	putative	EMR73267.1	Archaea
Thaumarchaeota_archaeon_SCGC_AB-539-E09	putative	EMR74324.1	Archaea
Thermacetogenium_phaeum_DSM_12270	corrinoid	WP_015049741.1	Bacteria
Thermacetogenium_phaeum_DSM_12270	corrinoid	WP_015051791.1	Bacteria
Thermacetogenium_phaeum	corrinoid	WP_015051800.1	Bacteria
Thermacetogenium_phaeum	Trimethylamine	KUK36621.1	Bacteria
Thermincola_ferriacetica	methanol-5-hydroxybenzimidazolylcobamide	WP_052218277.1	Bacteria
Thermincola_potens_JR	methanol-cobalamin	WP_013119873.1	Bacteria
Thermocladium_sp_ECH_B	hypothetical	KUO92439.1	Archaea
Thermofilum_carboxyditrophus_1505	methyltransferase	WP_052886742.1	Archaea
Thermofilum_pendens_Hrk_5	corrinoid	WP_011752875.1	Archaea
Thermofilum_pendens_Hrk_5	corrinoid	WP_011753130.1	Archaea
Thermofilum_sp_1807-2	hypothetical	WP_052884902.1	Archaea
Thermofilum_sp_1807-2	methyltransferase	AKG39330.1	Archaea
Thermofilum_sp_1807-2	methyltransferase	WP_052883866.1	Archaea
Thermofilum_sp_1910b	corrinoid	WP_020963330.1	Archaea
Thermofilum_sp_1910b	hypothetical	WP_020963087.1	Archaea

Continued on next page

Table 3-4 – continued from previous page

TaxonName	ProteinDescription	ProteinID	Taxonomy
Thermogladius_cellulolyticus_1633	corrinoid	WP_014737366.1	Archaea
Thermoplasmatales_archaeon_BRNA1	cobalamin-binding	WP_015492601.1	Archaea
Thermoplasmatales_archaeon_BRNA1	dimethylamine	WP_015491826.1	Archaea
Thermoplasmatales_archaeon_BRNA1	dimethylamine	WP_015492610.1	Archaea
Thermoplasmatales_archaeon_BRNA1	trimethylamine	WP_015492607.1	Archaea
Thermoplasmatales_archaeon_DG-70-1	methyltransferase	KYK32642.1	Archaea
Thermoplasmatales_archaeon_DG-70	methyltransferase	KYK38702.1	Archaea
Thermosediminibacter_oceani_DSM_16646	corrinoid	WP_013275879.1	Bacteria
Treponema_azotonutricium_ZAS-9	cobalamin-binding	WP_015710346.1	Bacteria
Treponema_primitia_ZAS-2	cobalamin-binding	WP_015709282.1	Bacteria
uncultured_archaeon	conserved	CBH38070.1	Archaea
uncultured_marine_thaumarchaeote_KM3_42_E08	methionine	AIF10042.1	Archaea
uncultured_marine_thaumarchaeote_KM3_70_E10	methionine	AIF15565.1	Archaea

CHAPTER 4

CHITINASES: A STANDARD CANDLE FOR DATING MICROBIAL LINEAGES

4.1 Abstract

Establishing the divergence times of groups of organisms is a major goal of evolutionary biology. This is especially challenging for microbial lineages, due to the near-absence of preserved physical evidence (diagnostic body fossils or geochemical biomarkers). Horizontal gene transfer (HGT), can serve as a temporal scaffold between microbial groups and other fossil-calibrated clades, potentially improving divergence time estimates for microbial lineages. Specifically, HGT to or from organisms with fossil-calibrated age estimates can propagate these constraints to additional groups that lack fossils. While HGT is common between lineages, only a small subset of HGT events are potentially informative for dating microbial groups. Constrained by published fossil-calibrated studies of fungal evolution, molecular clock analyses show that multiple clades of bacteria likely acquired chitinase homologs via HGT during the very late Neoproterozoic into the early Paleozoic. These results also show that, following these HGT events, recipient terrestrial bacterial clades likely diversified $\sim 400\text{-}500$ Ma, consistent with established timescales of arthropod and plant terrestrialization. We conclude that these age estimates are broadly consistent with the dispersal of chitinase genes throughout the microbial world being in direct response to the evolution and ecological expansion of detrital-chitin producing groups. The convergence of these multiple lines of evidence demonstrates the utility of HGT-based dating methods in microbial evolution. The pattern of inheritance of chitinase genes in multiple terrestrial bacterial lineages via HGT processes suggests chitinases can serve as a "standard candle" for dating microbial lineages.

4.2 Background

Dating when new metabolisms evolved and when major clades of Bacteria arose, particularly on the order of hundreds of millions of years, remain key challenges in biology (Dos Reis et al., 2015). Despite progress in understanding the molecular record of extant bacterial genomes, the timing of the evolution of major clades of Bacteria is especially challenging to resolve due to complex gene histories and a lack of clear phenotypic traits that can be correlated with a diagnostic fossil record. In the near-absence of physical (geochemical or fossil) records of microbial evolution, it is difficult to determine and date the evolutionary history of bacterial lineages.

Leveraging the information contained horizontal gene transfer (HGT) events can significantly improve estimates of the timing of events within microbial evolution (Wolfe and Fournier, 2018; Davin et al., 2017; Gold et al., 2017; Soucy et al., 2015; Dos Reis, 2018; Magnabosco et al., 2018). Vertical inheritance passes genetic information from parent to offspring, but HGT passes genetic information between organismal lineages, across all de-

degrees of evolutionary distance. This can be particularly useful for molecular clock dating, as HGTs establish cross-cutting relationships between lineages and serve as a "temporal scaffold" upon which fossil calibrations or other date information from even distantly related taxa may be placed (Dos Reis, 2018). While HGT is a major process in microbial evolution (Soucy et al., 2015; Andam et al., 2010), HGT events between microbes and eukaryotes with a fossil record are less frequent (Husnik and McCutcheon, 2018). Furthermore, the donor-recipient relationships are often difficult to infer for many HGT histories, due to multiple HGT events and gene losses, or lack of a strong phylogenetic signal (Fournier et al., 2009). The function of a gene is not necessarily relevant to its utility in propagating time constraints (*e.g.*, Wolfe and Fournier, 2018); however, in some cases, this gene function may be additionally informative, and provide a source of independent validation of age estimates. This is the case, for example, if the protein encoded by the transferred gene is specific for a substrate that can, itself, be temporally constrained. Given all of these criteria, a very small number of HGT events may be especially valuable for dating microbial lineages; these "index transfers" (Wolfe and Fournier, 2018) can be even more valuable if multiple HGT recipients are present, closely correlating the ages of the recipients in time, a "standard candle" (a term used in astronomy to describe an object with known luminosity used to infer the cosmic distances to other objects of interest) (Colgate, 1979).

4.2.1 Environmental Distribution of Chitin

Chitin is one of the most abundant structural polysaccharides in nature (Bai et al., 2016; Talamantes et al., 2016), and chitin degradation by chitinases is a critical process in the biogeochemical cycling of carbon and nitrogen in terrestrial and aquatic ecosystems (Bai et al., 2016). Biogenic sources of environmental chitin include fungi (Funkhouser and Aronson, 2007), arthropods (Merzendorfer and Zimoch, 2003), marine invertebrates (Yoshioka et al., 2017), bacteria (Chang and Stergiopoulos, 2015), and corals (Bo et al., 2012). However, chitin is produced mainly by arthropods and fungi, and is thought to be present in higher abundance today in the terrestrial, rather than marine, system (Talamantes et al., 2016) following the terrestrialization of arthropods, sometime after the Cambrian (Schwentner et al., 2017). In modern aquatic systems, Arthropods are the dominant chitin-producing organisms. The annual global chitin sourced from Arthropods is estimated to be 2.8×10^7 Mg yr⁻¹ for freshwater ecosystems to 1.3×10^9 Mg yr⁻¹ for marine ecosystems (Cauchie, 2002). The majority of chitin in terrestrial ecosystems is produced by fungi (Smrř and Čatská, 2010) largely due to their greater contribution of biomass to the soil environment (Holtkamp et al., 2008). While global estimates for the contribution of arthropod biomass, and thus chitin, to the environment over time are lacking, they nonetheless make up the largest pool of animal biomass today (Bar-On et al., 2018).

4.2.2 Chitin Production and the Evolution of Fungi

The evolution of chitin producers, fungi and marine and terrestrial arthropods, is anchored to the fossil record through diagnostic morphological characters (James et al., 2006; Berbee and Taylor, 2010; Berbee et al., 2017; Strullu-Derrien et al., 2018; Taylor and Berbee, 2006). Cryptomycota form the most deeply branching fungal clade, and contain the most deeply branching chitinous Fungi (*e.g.*, Rozella) (Berbee et al., 2017; Strullu-Derrien et al., 2018). Fossil-calibrated molecular clock studies generally agree that early Fungi diverged around 1125-900 Ma (Parfrey et al., 2011; Sharpe et al., 2015). Fossil and molecular clock evidence also indicates that divergence of Ascomycota and Basidiomycota within the major fungal group Dikarya occurred around 830-518 Ma (Berbee et al., 2017; Strullu-Derrien et al., 2018). Fossil-calibrated molecular clock studies generally agree that early Fungi diverged around 1125-900 Ma (Parfrey et al., 2011; Sharpe et al., 2015). Fossil and molecular clock evidence also indicates that divergence of Ascomycota and Basidiomycota within the major fungal group Dikarya occurred around 830-518 Ma (Berbee et al., 2017) with a fossil minima around 405 Ma (Berbee and Taylor, 2010; Floudas et al., 2012; Wolfe et al., 2016). Secondary calibrations from molecular clock studies suggest that crown Ascomycota diversified 715-408 Ma (Prieto and Wedin, 2013) and crown Basidiomycota diversified 655-400 Ma (Floudas et al., 2012). Therefore, studies of fungal evolution can provide strong secondary calibrations for dating chitinase gene trees.

Based on fossil and molecular clock dating methods, marine crown-group euarthropods appeared around 521-514 Ma, shortly after the start of the Cambrian, and radiated into the lower and middle Cambrian (Daley et al., 2018; Wolfe et al., 2016). Molecular clock and fossil evidence suggests that terrestrialization of major arthropod groups occurred from the Cambrian into the Silurian (Lozano-Fernandez et al., 2016). The oldest terrestrial myriapod body fossil (the oldest undisputedly terrestrial animal) is the 426 Ma millipede *Pneumodesmus newmani*, from the Silurian of Scotland (Wilson and Anderson, 2004). However, the radiation of terrestrial arthropods (including insects) likely continued into the Devonian (Oakley et al., 2013; Schwentner et al., 2017; Glenner et al., 2006).

4.2.3 The Evolution of Chitinase Gene Families

Chitinases are proteins that catalyze the breakdown of glycosidic linkages in polymers of chitin (Funkhouser and Aronson, 2007). Chitinases are a type of glycoside hydrolase (GH) specific to chitin (Talamantes et al., 2016; Berlemont and Martiny, 2015). There are two main families of chitinases: glycoside hydrolase family 18 (GH18) and glycoside hydrolase 19 (GH19) (Funkhouser and Aronson, 2007). GH18 chitinases are distributed across the three domains of life (Chang and Stergiopoulos, 2015; Funkhouser and Aronson, 2007), whereas GH19 chitinases are restricted mostly to plants and are rarely associated with

bacteria (Chang and Stergiopoulos, 2015). In one well-studied bacterial model organism, *Streptomyces*, there were ten genes associated with the GH18 family of chitinases (homologs chiA-E, and H- L) and two genes associated with GH19 (chiF, G) (Ubhayasekera and Karlsson, 2012). It has been suggested that some of these genes may have evolved under selective pressures related to the host environment or to the presence and proximity to other organisms, which may have even precipitated HGT events (Ubhayasekera and Karlsson, 2012; Ihrmark et al., 2010; Mamarabadi et al., 2008). Myxobacterial chitinases have been hypothesized to have evolved via HGT (Sharma and Subramanian, 2017), and other bacterial lineages within Actinobacteria are hypothesized to have co-opted a fungal chitinase for self-defense (Ubhayasekera and Karlsson, 2012). Because of the specific associations between substrate and gene, it stands to reason that there may be an evolutionary link between the major producers of environmental chitin (fungi, bacteria, and arthropods) and the genes that break it down. It has been shown that some bacterial chitin degradation systems are even adapted to the environments (aquatic vs. terrestrial) and most abundant chitin producers (exoskeletons of crustaceans vs. fungal cell walls) that they encounter (Bai et al., 2016). Nonetheless, it remains to be tested whether chitinase genes also reflect widespread environmental adaptations over geological time.

It has been shown that chitinases may retain a molecular record of evolutionary events hundreds of millions of years ago (Emerling et al., 2018). Moreover, while some of the phylogenetic distribution of these genes may indicate a pattern of vertical inheritance, other chitinase genes may have evolved via horizontal gene transfer (Ubhayasekera and Karlsson, 2012). For these reasons, and the criteria described above, chitinase genes are an attractive potential source of temporal information for microbial evolution. Therefore, we sought to test the hypothesis that specific bacterial chitinases evolved via HGT, and if so, if these HGT events could be leveraged to propagate known fossil calibrations between donor and recipient lineages. Bacterial chitinases are especially useful because they metabolize chitin, a specific biopolymer only produced in abundance by arthropods and fungi, two groups with fossil records and thus likely age estimates much more precise than those of most microbial groups. Previous work has also suggested that some chitinases are distributed between the domains of life via HGT, for example, postulating that some chitinase genes were transferred from plants to Actinobacteria and then to arthropods (Lacombe-Harvey et al., 2018). However, the evolutionary history of the many disparate chitinase gene families in microbes has not been fully investigated.

4.2.4 Bayesian Molecular Dating

Fossil-calibrated molecular clock models are applied to estimate divergence times of organisms (e.g. (Schirmer et al., 2015; Donoghue and Yang, 2016)). Many molecular clock model parameters have only been recently developed, and few have been applied to microbes

with divergence time estimates that span geologic time or have undergone rampant horizontal gene transfer events (e.g. (Dos Reis, 2018; Louca et al., 2018)). For a more detailed review of these parameters and challenges see, for example, (Hillis et al., 2005; Drummond et al., 2006; Donoghue and Benton, 2007; Edwards, 2009; Heled and Drummond, 2015; O'Reilly et al., 2015; Donoghue and Yang, 2016; Bromham et al., 2018; Knoll, 2017). The issues inherent to assessing microbial evolution present a challenge for this work, but also an opportunity to explicitly test these model parameters and assumptions in order to determine those that are valid for this specific set of evolutionary conditions.

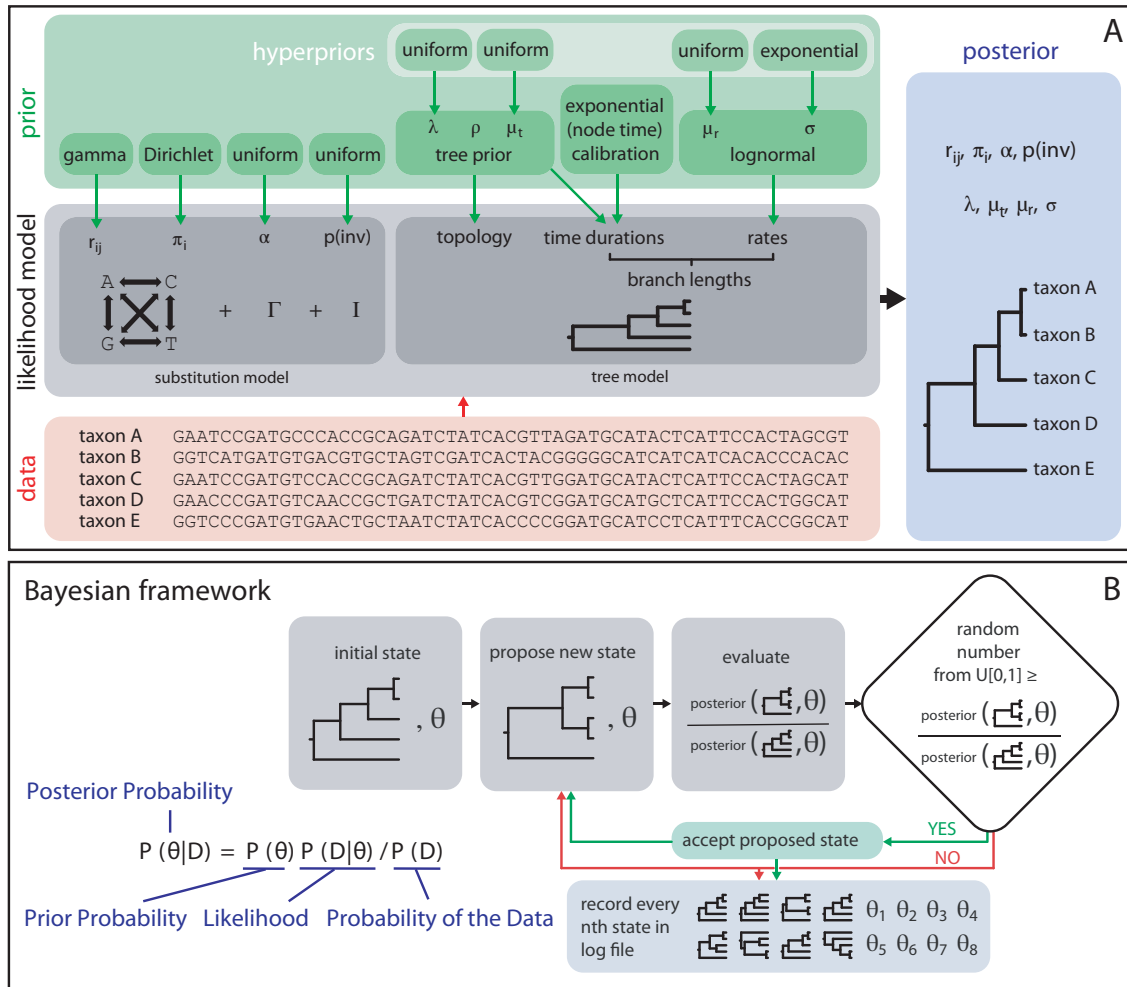


Figure 4-1: Bayesian Molecular Dating Parameters. Modified from (Bromham et al., 2018). The upper panel (A) illustrates the components of these analyses and the lower panel (B) depicts a flowchart of how the posterior is obtained next to Bayes's theorem.

Molecular clock dating is based on a Bayesian framework, reviewed in greater detail by others (Dos Reis et al., 2015; Bromham et al., 2018; Nascimento et al., 2017). There are a few major components used to determine posterior probabilities or date distributions such as data selection, calibrations, the molecular clock model, the tree process prior, and the rate distribution model. Figure 4-1 illustrates these concepts. These parameters will be tested in subsequent sections of this chapter, but are first defined here. The sequence data assessed in this work are the chitinase genes present in bacterial and eukaryotic lineages. Tree process priors include birth-death and uniform. Rate distribution models include lognormal autocorrelated and uncorrelated gamma.

Tree priors tested include the uniform prior and the birth-death process. The uniform prior considers every possible topology to be equal and favors divergences that are evenly spaced across the tree from the root to tip (Huelsenbeck et al., 2002; Lepage et al., 2007). The birth-death model is defined by speciation ("birth") and extinction ("death"). In contrast to the uniform prior, this tree process ascribes more weight to tree topologies with certain branching patterns (Rannala and Yang, 1996). The birth-death process generally biases the model such that deeper branches are longer and the more shallow branches are shorter, because it is assumed the "older" lineages more often end in extinction (Bromham et al., 2018). Biases such as this can have large effects on the posterior age estimates and inappropriate model selection can result in less precise dates.

All models in this study assume a relaxed molecular clock model for a prior on the branch rate. However, two relaxed clock models for the branch rates are assessed: autocorrelated and uncorrelated. Uncorrelated clocks make no assumption that branches next to each other on the tree should share similar rates. In other words, the rate on each branch of the tree is independent. Conversely, autocorrelated clocks assume that more closely related branches on the tree should also have more similar rates (Thorne et al., 1998; Drummond et al., 2006; Lepage et al., 2007; Ho and Duchêne, 2014). The assumption that neighboring branches should share more similar rates makes sense when we consider that the evolution of genetic information between related lineages is often affected by many of the same processes that affect the rates of evolution (e.g. environment, population) (Bromham et al., 2018). Biological events such as horizontal gene transfer may invalidate model assumptions, but the mechanisms of rate variation and quantifying the relative importance of various biological events is still debated (dos Reis et al., 2016). Choosing between these models is a matter of ongoing debate in the field, and is often dependent on the data set (Lepage et al., 2007; Lartillot et al., 2009; Bromham et al., 2018). Thus, we detail the effects of model selection in our analyses.

4.2.5 Summary of Hypotheses and Objectives

The primary objective of this work is to test whether fossil-calibrated age estimates within fungi can be propagated to bacterial lineages through the use of HGT events between these lineages under different model assumptions. Secondly, we seek to understand possible ecological implications of the evolution of chitinases in fungi and bacteria. If bacterial chitinase genes were acquired in response to environmental chitin availability, then arthropod evolutionary history provides a prediction for the timing of these events within bacterial lineages. We hypothesize that terrestrial bacterial chitinases diversified from the Cambrian into the Devonian following the distribution of environmental chitin. We independently date chitinase evolution in microbial lineages by first testing and then applying molecular clock models to chitinase gene trees, constrained by fungal date calibrations tethered via HGT. We show that certain model parameters seem to outperform others. Moreover, our posterior date distributions for bacterial lineages support the utility of HGT-propagated fossil calibrations in accurately estimating the ages of microbial lineages as an avenue for future work.

4.3 Methods

4.3.1 Taxon Sampling

We queried The National Center for Biotechnology Information (NCBI) nonredundant (nr) database using the protein Basic Local Alignment Search Tool (BLASTp) for sequences homologous to the *Myxococcus fulvus* ChiD protein (WP_046715376.1). Complete protein sequences of the top 5000 hits from NCBI were downloaded (E-value < 10⁻⁵). Sequences were subsampled to include one taxon per annotated species. We further used BLASTp to more exhaustively identify potential homologs within Fungi, repeating this method for specific searches within Ascomycota, Basidiomycota, and more deeply-rooting Fungi (*e.g.*, Blastocladiomycota, Chytridiomycota, Zoopagomycota, and Mucorales).

4.3.2 Sequences and Alignments

Sequences were aligned using MUSCLE (Edgar, 2004). Poorly aligning regions were identified via manual inspection and removed using Jalview (Waterhouse et al., 2009). Sequences were manually edited to correct obvious misalignments (*Fonsecaea multimorphosa* and *Phialophora americana* (sites 2390-2470), and removed the misaligned C-terminal region from *Phelbia centrifuga*, and the misaligned C-terminal regions from *Rhizopus*, *Mucor*, *Syncephalastrum*, *Absidia*, and *Lietheimia* (sites 2393 onward).

A profile alignment of bacterial and fungal sequences was made (Edgar, 2004). This revealed a highly conserved alignment region shared across bacterial and fungal sequences

(sites 1844-2470) and another well-aligned N-terminal region conserved across Bacteria, but absent or poorly aligned in Fungi. In order to maximize the sequence information used for phylogenetic reconstruction and molecular clocks without introducing misalignments between bacterial and fungal sequences, a composite alignment was generated. This involved concatenating the conserved region for both Fungi and Bacteria with the N-terminal region aligned for just Bacteria. From this alignment, a single gene tree could be generated for determining the relationship between Fungi and Bacteria, but also maximally resolving splits within the bacterial tree.

4.3.3 Phylogenetic Analyses

Gene Tree

The gene tree were inferred using RaxML v1.8.9 using the PROTGAMMALGF substitution model (Stamatakis, 2006) as fit by PROTTEST (Darriba et al., 2017), and 100 bootstrap replicates. The resulting tree showed relationships between fungal taxa congruent with published phylogenies (Taylor and Berbee, 2006; Berbee et al., 2017; Strullu-Derrien et al., 2018; Berbee and Taylor, 2010; James et al., 2006). We rooted the gene tree on the branch leading to *Rozella*, which is considered to be either sister to the most deeply-rooting fungal clades, or a member of Chytridiomycota, one of the most deeply-rooting Fungi (Berbee et al., 2017). This resulting rooting placed bacterial chitinases as a clade diverging within crown Fungi, polarizing the origin of the bacterial homologs as originating via an HGT from a fungal donor.

Divergence Time Estimation

Divergence times were estimated using PhyloBayes v3.3 under the CAT20 set of substitution models (Lartillot et al., 2009). Divergence time estimates were generated under several sets of model priors. Specific model parameters are described in Tables 4-1 and 4-2. After chain convergence (effective size >50 , variable discrepancies <0.30), trees and posterior probability support values were generated from completed chains after the initial 20% of sampled generations were discarded as burn-in.

Date Constraints

Secondary calibrations were applied to the divergence times of major fungal groups within the gene tree. For all analyses, we applied a root prior and one internal date constraint to the split of Ascomycota and Basidiomycota consistent with reported molecular clock and fossil evidence within Fungi (James et al., 2006; Taylor and Berbee, 2006; Berbee et al., 2017; Strullu-Derrien et al., 2018; Parfrey et al., 2011; Sharpe et al., 2015; Floudas et al., 2012; Prieto and Wedin, 2013; Wolfe et al., 2016; Berbee and Taylor, 2010). In order to

avoid false precision, uniform priors were used in both cases, 1145-739 Ma for the fungal root (Sharpe et al., 2015) and 830-518 Ma for the Ascomycota-Basidiomycota split (Floudas et al., 2012). We also tested the addition of secondary calibrations on the nodes leading to the Ascomycota (715-408 Ma) and Basidiomycota (655-400 Ma) clades (Prieto and Wedin, 2013; Floudas et al., 2012). Finally we tested the application of a fossil minima to the split on Ascomycota and Basidiomycota (830-405 Ma) (Berbee and Taylor, 2010; Floudas et al., 2012; Wolfe et al., 2016). All calibration structures are listed in Table 4-2.

4.4 Results

4.4.1 Phylogeny of ChiD and ChiC homologs

Figure 4-2 illustrates the relationship between taxa in this study (Table E-1) as a maximum-likelihood gene tree generated with RAXML. The tree is rooted with the most deeply-branching fungal taxon, *Rozella* (Cryptomycota). The group of deeply-rooting Fungi include members of Cryptomycota, Blastocladiomycota, Chytridiomycota, Blastocladiomycota, Chytridiomycota, Mucormycotina, and Zoopagomycota (in order of branching from the root). Bootstrap supports are low for many bipartitions within this deeply-rooting group. Support for the bipartitions placing bacterial sequences within Fungi are higher (74, 71). Support for the monophyly Ascomycota and Basidiomycota is high (100). Support is also high for the monophyly of bacterial sequences (99). While the deeper branches in the fungal tree have weak bootstrap support, the relatively short branches relating these groups and the lack of any calibrations sensitive to their specific crown-group topology suggest the observed phylogenetic uncertainty has little impact on divergence times for more distal clades within the tree.

Within Bacteria are the generally well-supported and often monophyletic Bacterial clades including Betaproteobacteria, Deinococcus, Actinobacteria, Bacteroidetes, Firmicutes, and Deltaproteobacteria. Gammaproteobacteria are polyphyletic, including Vibrionales, Xanthomonadales, and one Gammaproteobacteria Taxon in Actinobacteria (*Cellvibrio*, WP_049631752.1, a cellulolytic bacterium in the order Pseudomonadales (Mergaert et al., 2003), suggesting multiple independent acquisitions of ChiD Actinobacteria (bootstrap support 99), Bacteroidetes (bootstrap support 42), Firmicutes (bootstrap support 79), and Deltaproteobacteria (bootstrap support 88) are also monophyletic. Deltaproteobacteria sit on a reticulating branch within Firmicutes.

4.4.2 Divergence Time Estimates of Bacterial Chitinases

Divergence time estimates were tested under several models, evaluating the impact of, taxon sampling (inclusion or exclusion of bacterial sequences) effective priors (inclusion or exclu-

sion of sequence data, tree priors (uniform vs. birth-death), and relaxed clock models (autocorrelated lognormal vs. uncorrelated gamma rate distributions). Our preferred model is uncorrelated gamma distribution under a uniform prior with calibrations on the root (1145-738), Dikarya (830-518) and crown Ascomycota (715-408) and crown Basidiomycota (655-400).

Few published age estimates exist for the bacterial clades present in our tree. For example, based on the chitinase HGT from a time-calibrated Fungi tree, the effective prior for crown-group Vibrionales is ~ 300 Ma, with an uncertainty spanning from ~ 200 -400 Ma. The only other published divergence time estimate for Vibrionales (the last common ancestor of *Vibrio* and *Photobacterium*) was an uncalibrated RelTime clock on 16S and other protein datasets (Marin et al., 2017). The result for this clade was 124 Ma. Based on the chitinase HGT from a time-calibrated Fungi tree with a uniform prior and uncorrelated gamma clock model, the posterior age estimate for crown-group Vibrionales is ~ 188 Ma with an uncertainty spanning ~ 278 -113 Ma. This indicates that our dataset propagating fungal fossil calibrations via a relatively short chitinase gene (instead of 16S sequence data) may nevertheless provide reasonable age ranges for bacterial taxa.

The chronogram depicted in Figure 4-3 shows that bacterial chitinases have a common ancestor ~ 780 Ma (Node 3, Table 4-3) and were acquired from fungi prior to the evolution of marine arthropods in the Cambrian. Subsequent HGT events between bacterial groups distributed this gene, with the major bacterial clades in the tree acquiring chitinase ~ 505 -188 Ma. This age range is consistent with the ecological and taxonomic dispersal of bacterial chitinases being correlated with the evolution and diversification of crown group euarthropods around 521-514 (Wolfe et al., 2016; Daley et al., 2018). Interestingly, four major clades of terrestrial Bacteria in the tree, Gammaproteobacteria (Xanthamonadales), Betaproteobacteria, Actinobacteria, Firmicutes, all diversify ~ 408 -365 Ma, temporally consistent with the terrestrialization of marine arthropod groups most conservatively estimated to have occurred 541 - 359 Ma (Lozano-Fernandez et al., 2016; Schwentner et al., 2017; Oakley et al., 2013) (Figure 4-2). This timing is also consistent with the early terrestrialization of land plants (middle Cambrian - Early Ordovician) and vascular plants (Late Ordovician - Silurian) (Morris et al., 2018), and alternatively, may represent the establishment of plant-degrading Fungi in soils around 300 Ma (Floudas et al., 2012).

4.4.3 Testing Molecular Clock Models

Molecular clock model as listed in Table 4-1 were tested to assess model parameters sensitivities. The results for Model 6 (selected for further analysis) are presented in Table 4-3. The results of all model outputs are listed in Table E-2. An analysis of these models is presented in Table 4-4 and further elaborated upon in the following sections. Table 4-4 illustrates the models excluding secondary calibrations on Ascomycota and Basidiomycota

crown groups that recover the expected age ranges for these nodes in the literature. For this analysis, Calibrations 1 and 3 were used, as these do not impose dates on crown Ascomycota or Basidiomycota clades, enabling comparison between estimated and expected model output for these clades. Table 4-4 shows that the 95% CI ages fall within expected ranges for the uniform prior and uncorrelated gamma relaxed clock model for Ascomycota under Calibrations 1 and 3. The model ages also fall within expected ranges the uniform prior and uncorrelated gamma clock model for Basidiomycota Calibration 1; uniform prior and lognormal autocorrelated clock model for Ascomycota, Calibration 3; and birth-death prior and uniform gamma distributed model for Basidiomycota, Calibration 3. Mean ages for the birth-death prior and uncorrelated gamma model and for the uniform prior with lognormal model fall outside of expected age ranges under Calibrations 1 and 3 for Ascomycota and under Calibration 1 for Basidiomycota.

4.4.4 Impact of the Tree Process Prior and Rate Distribution Model

The effects of the tree process prior (birth-death vs. uniform) and the rate distribution model (lognormal correlated vs. uncorrelated gamma) were evaluated (Table 4-4, Table E-2). Prior and posterior age estimates for the chitinase tree using a uniform vs. birth-death prior and lognormal vs. gamma distribution result in different date distributions across nodes, in both bacterial and fungal groups. Across the Bacterial nodes, the uniform prior with lognormal autocorrelated clock model corresponded to the oldest date estimates across nodes, followed by the uniform prior and uncorrelated gamma model, birth-death prior and lognormal autocorrelated model, and finally the youngest birth-death prior and uncorrelated gamma relaxed clock model (Figure 4-4). The birth-death prior resulted in the youngest age estimates as compared to the uniform prior (Figure 4-4). The same pattern holds for the Ascomycota and Basidiomycota within the Fungal nodes. However, a slightly different result is observed for the deeply-rooting Fungal nodes (root, Fungi, and Dikarya). For these Fungal nodes, the opposite pattern is seen with the oldest date distributions resulting from the birth-death prior and uncorrelated gamma clock model, followed by the birth-death prior and lognormal autocorrelated model, the uniform prior and uncorrelated gamma model, and finally the youngest uniform prior and lognormal autocorrelated clock model (Figure 4-4). This empirical internal control on predicting fungal age estimates for nodes that have had their calibrations removed suggests that the uniform tree process and uncorrelated gamma rate distribution provide the most accurate age estimates for this gene family.

This model selection is also theoretically justifiable. A birth-death prior is a tree process prior that assumes a tree generated by speciation and extinction events across a lineage (Heled and Drummond, 2015). This assumption is violated for trees that include HGT events, especially if several such events are present. Birth-death priors are therefore not appropriate for gene trees that show histories of extensive HGT, since the underlying as-

sumption, that nodes are distributed across a continuity of lineage speciation and extinction, is invalid. This is especially true for HGTs between prokaryotes and eukaryotes, which likely have very different patterns of speciation and extinction occurring over very different timescales, and also very different sampling densities. The chitinase tree is an especially good test of these hypotheses, as in this dataset we infer multiple HGTs between Bacteria after a primary HGT from Fungi. There are many nodes that are clearly not the consequence of birth-death processes. In fact, the ecological dispersal of genes via HGT should be expected to locally increase node densities in the tree entirely independent of any underlying assumptions of speciation or extinction. In the absence of a different model sensitive to nodes mapping to be transfer vs. speciation, it is important to avoid assumptions made in the birth-death model. In addition, for many of the bacterial nodes, the uniform tree process prior results in broader prior ages than the birth-death prior. Therefore, the violation of the assumptions of a birth-death process in the bacterial chitinase tree may result in overly narrow priors that are too informative. Additionally, autocorrelated rate distribution models generally underperform for large evolutionary distances (Drummond et al., 2006), and inspection of the gene tree does not readily reveal any lineage-specific branch length effects that suggest rate biases that would be poorly accounted for under an uncorrelated model.

4.4.5 Impact of Taxon Sampling and Fungal Divergence Times

The impact of taxon sampling was evaluated (Tables E-1, E-2). Within Fungi, the chitinase gene appears to follow a history of vertical descent, and therefore better modeled under a birth-death tree process prior. Therefore, one test of the appropriateness of a birth-death process prior is if the presence of bacterial sequences within the tree impacts the effective prior ages within Fungi. Ascomycota and Basidiomycota groups each have prior ages ~ 100 Ma younger under the birth-death model when Bacteria are removed. Under the uniform model, Ascomycota is the same age whether or not Bacteria are included, while Basidiomycota is also ~ 100 Ma younger. In general, the birth-death model gives much younger prior ages (Figure 4-5), ~ 150 Ma for Bacteria and Basidiomycota, whether or not Bacteria are in the tree, and ~ 150 Ma for Ascomycota in the presence of Bacteria, and ~ 250 Ma in the absence of Bacteria. Ascomycota and Basidiomycota crown group age priors are very sensitive to the tree process prior. Therefore, we chose to use additional secondary calibrations within Dikarya to aid in guiding the prior on the Ascomycota and Basidiomycota nodes.

4.4.6 Impact of Calibrations

In general, the date distributions across all nodes do not appear to be very sensitive to the calibrations applied under the uniform distribution and uncorrelated gamma relaxed

clock model. Because the calibrations are all roughly in the same range, it appears that all calibration models lead to similar date distributions (Figure 4-5). However, Calibration 2 (calibrations on the root, spit of AB, and A and B crown lineages, not including the fossil minima) lead to slightly more precise peaks (Figure 4-5). There are two potential problems with using single gene alignments to generate a posterior age estimate for an HGT: a single gene has limited rate information from aligned sites for an informative molecular clock, and, if HGT increases the rate of evolution along reticulate branches due to genes evolving faster once in a recipient genome, then posteriors will bias towards under-estimating the ages of these groups. Therefore, we wanted to assess whether younger posterior dates generated by the birth-death prior as compared to the uniform prior were due to the long branch separating Bacteria from Fungi in the tree. It is possible that this long branch may either be representative of a longer time interval (and thus younger crown ages) or of a faster evolutionary rate (and thus older crown ages). The maximum likelihood tree (Figure E-1) illustrates that when rooted, the Ascomycota and Basidiomycota actually have slightly longer distances to the root, suggesting that the relative rates of evolution in this gene tree are not accelerated in the bacterial group. Consequently, the limited sequence information contained in this dataset may be used to calculate posterior age estimates that are unlikely to be biased by HGT-induced rate effects. Including additional internal constraints on the fungal clades push the priors under the uniform and birth-death models closer together for Bacterial nodes. These additional secondary calibrations are thus important for constraining the tree process prior, and this type of approach may be important for using single gene HGTs to improve age estimates in general.

4.4.7 Informativeness of Sequence Data

We assessed the informative of the sequence data by running Phylobayes under the prior (effective prior, including calibrations (Fig. E-2)). Posterior age distributions for bacterial chitinase nodes substantially differed from prior age distributions, showing that sequence data is meaningfully informing age estimates via the relaxed molecular clock (Figure 4-6).

4.5 Discussion

4.5.1 Fungal Origin and Distribution of Bacterial Chitinases

The gene tree topology for ChiC/D and its inferred rooting within Fungi show that bacterial chitinase was acquired via HGT from a fungal donor lineage. By including secondary age calibrations on nodes within Fungi, molecular clock estimates show that this gene was acquired by Bacteria by 605 Ma (range of 655-566 Ma), slightly predating estimates for the evolution of crown arthropods (marine euarthropods) (Daley et al., 2018). The earliest

Fungi were also marine, and likely colonized land by moving from marine to terrestrial environments (Lozano-Fernandez et al., 2016; Berbee et al., 2017). This supports the hypothesis that some bacterial chitinases evolved from a marine fungal ancestor.

The HGT between fungi and bacteria also seems plausible from environmental and mechanistic perspectives. Not only do bacteria and fungi occupy similar environments, but closely-related fungal relatives also include both phagotrophic Mucormycotina and pathogenic Entomophthorales (James et al., 2006), suggesting plausible scenarios for this transfer. This is also consistent with evidence that other bacterial chitinases within the GH18 family (*e.g.*, ChiJ) also evolved via HGT from Fungi (Ubhayasekera and Karlsson, 2012) although this particular HGT event would have been more recent than the HGTs discussed here. Following the initial transfer into a bacterial lineage, bacterial groups look to have all acquired chitinases from one another via subsequent HGT events, although the donors of these HGTs can't be directly inferred from the tree topology, except in the case of Firmicutes to Deltaproteobacteria.

4.5.2 Importance of Chitinase Evolution for Dating Microbial Metabolisms

Nearly all bacterial groups lack fossil evidence that could potentially constrain crown-group clades. There are some fossil constraints within Cyanobacteria (Schirmermeister et al., 2015) and other bacterial lineages contain proxy eukaryote fossil calibrations, such as mitochondrial lineages within Alphaproteobacteria (Magnabosco et al., 2018). Nonetheless, taxa within major lineages such as Firmicutes, which are distant relatives to these better-calibrated groups, are difficult to date, and because they are so distant, calibrations for other regions of the tree, even if they exist, are essentially not informative.

Substrate-specific genes, such as chitinases, are generally valuable for placing older-bound ages on microbial lineages, in the narrow case if these genes are vertically inherited within a group, and in the broader case if their origin is deeper, with subsequent HGT. In addition to the divergence time estimates shown for bacterial lineages in Figure 4-3, the HGTs themselves can provide useful temporal constraints for subsequent analyses (Wolfe and Fournier, 2018). Our results suggest that there was likely a transfer of chitinase genes between Bacilli (Firmicutes) and Myxococcales (Deltaproteobacteria) (Figure 4-2). Because these clades are very distant on the species tree (Raymann et al., 2015) being able to link them via the HGT of this chitinase gene means that we can apply relative constraints between groups. Divergence time age estimates from this study can also be useful for future investigations. While a single gene, such as ChiD, contains little sequence data for informing posterior age distributions, fossil calibrated effective priors can be used as constraints for other molecular clock studies that use far larger sequence datasets that recover species tree relationships, but cannot be directly fossil calibrated.

In general, it appears that numerous clades of bacteria acquired chitinase genes dur-

ing the early Paleozoic, suggesting that their dispersal throughout the microbial world was in direct response to the evolutionary and ecological expansion of detrital-chitin producing arthropod groups, upon which the bacteria may have made their habitats and/or food sources. The timing of bacterial chitinase divergence also coincides with the expansion of terrestrial organisms including both land plants and arthropods. The temporal and environmental constraints on chitin as a substrate support our results. For example, a molecular clock under an autocorrelated log-normal rate distribution produces mean posterior age estimates for inheritance of chitinases within terrestrial bacterial clades ~ 550 -500 Ma, before the terrestrialization of either arthropods or plants. Mean age estimates from uncorrelated models are both substantially younger and more congruent with the terrestrialization fossil record, as well as distributed over a much narrower time interval, suggestive of a causal relationship (~ 450 -150 Ma) (Figure ??). While this is a reasonable inference, this does independently support the hypothesis, as well as validate the HGT dating method used here. Moreover, because the priors under a uniform tree process are not subject to violations exposed in the birth-death model, we do not see evidence of HGT-associated heterotachy within the bacterial recipient groups, and there is limited site information in the gene for estimating rates for a posterior. Posterior dates used as secondary constraints on future species may be able to increase precision (Yang and Rannala, 2005; Rannala and Yang, 1996).

4.5.3 Ecological Implications of Chitinase Evolution

In general, our results show that numerous clades of bacteria acquired chitinase genes during the early Paleozoic, suggesting that their dispersal throughout the microbial world was in direct response to the evolutionary and ecological expansion of detrital-chitin producing arthropod groups. The temporal and environmental constraints on chitin as a substrate are consistent with these results. The timing of chitinase genes holds ecological implications for the origin of environmental chitin production. It is uncertain how the primary origin of environmentally-relevant amounts of chitin has evolved through time; did this originate from fungal cell walls or detrital chitin from the molted exoskeletons of arthropods? The genetic record may aid in distinguishing these sources. Robust across our model parameters and assumptions, bacterial chitinase (ChiD) appears to have evolved from fungi, likely in response to the availability of chitin as a major structural component of fungal cell walls in the Proterozoic, prior to early arthropod evolution. Subsequent HGT and inheritance of chitinase within terrestrial bacterial clades appears to be a much more recent series of evolutionary events within the early Paleozoic, consistent with evidence for plant and arthropod terrestrialization during this time (Lozano-Fernandez et al., 2016; Kenrick et al., 2012). The taxonomic distribution of ChiD within marine microbial groups appears to be too sparse to infer the timing of their acquisition, or to polarize the deep HGT events

between microbial lineages, which, presumably, progressed from marine to terrestrial clades. The long reticulating branch leading from Fungi to the bacterial ChiD sequences suggests that the direct fungal donor clade is not represented in the current tree; this may be due to unsampled extant fungal diversity, or patterns of extinction among ancient marine fungal groups.

4.6 Conclusion

Bacterial chitinases appear to have diversified from acquisition, roughly 600 Ma, into the last 200 Ma of Earth history. This is consistent with the hypothesis that bacterial chitinases evolved in response to the seeding of marine and terrestrial environments with globally-significant amounts of chitin, first from Fungi, then later from marine and then terrestrial arthropods. There is later evidence of at least one major HGT event within bacterial lineages, from Firmicutes to Deltaproteobacteria. Although we only assessed one chitinase gene tree in this study, future work evaluating the phylogenetic distribution of other chitinase genes will be critical for quantifying chitinase evolution in marine and terrestrial environments to further test the hypothesis that the phylogenetic distribution of chitinase genes mirrors the evolution and terrestrialization of environmental chitin sources.

Further, we show the importance of prior choice, highlighting that this dataset, which includes at least one deeply-rooted HGT, violates the birth-death prior. Moreover, we argue for the use of a uniform prior, uncorrelated gamma multipliers model, and three internal secondary calibrations propagating fossil calibrations from within Fungi to Bacteria.

Finally, we suggest that our dataset does not demonstrate HGT-associated heterotachy. Thus, our fungal priors and perhaps even posterior bacterial date distributions, may be more broadly applicable for future molecular clock studies assessing the divergence times of these major clades of Bacteria.

4.7 Acknowledgements

Thank you to Abby Caron for help with many of the pipelines and scripts that made this work so efficient and productive. Thank you to Jo Wolfe, Cara Magnabosco, and Greg Fournier for constructive feedback and discussions.

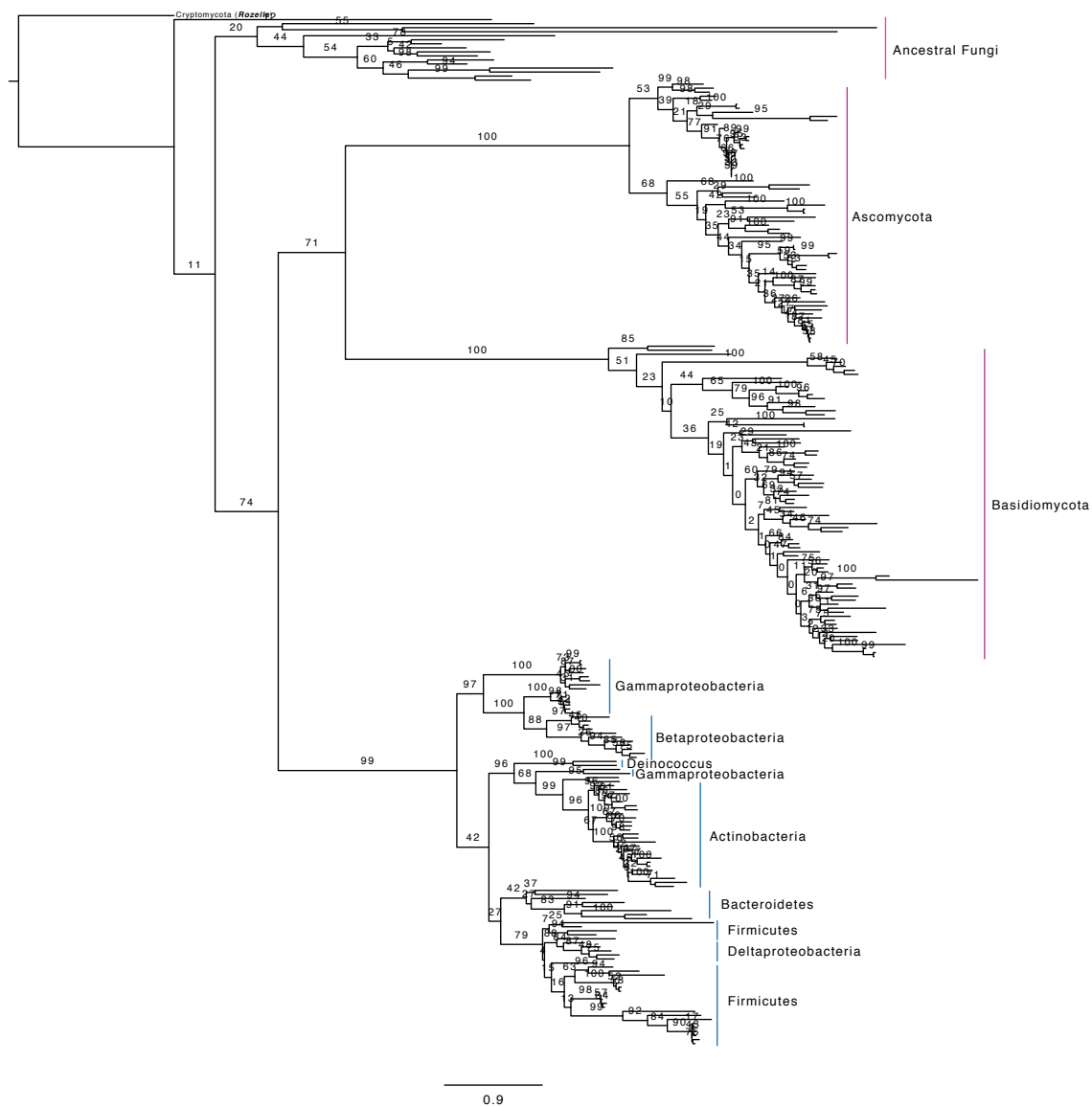


Figure 4-2: Chitinase Gene Tree. Maximum likelihood gene tree (RaxML) illustrating the relationship between fungal and bacterial taxa.

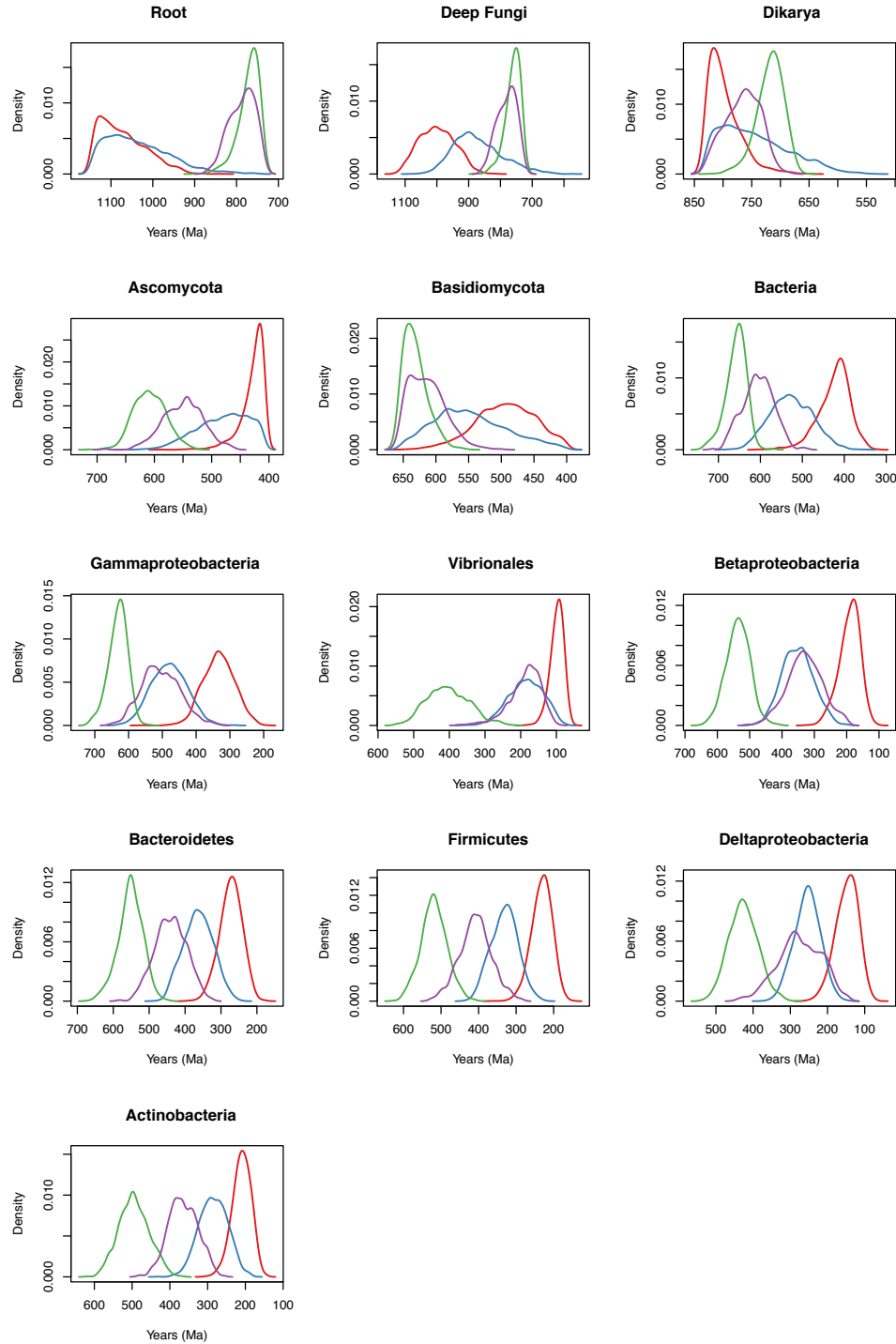


Figure 4-4: Posterior chitinase date distribution across nodes with varied model selection. These correspond to Models 2, 6, 10, and 14 and Calibration 2

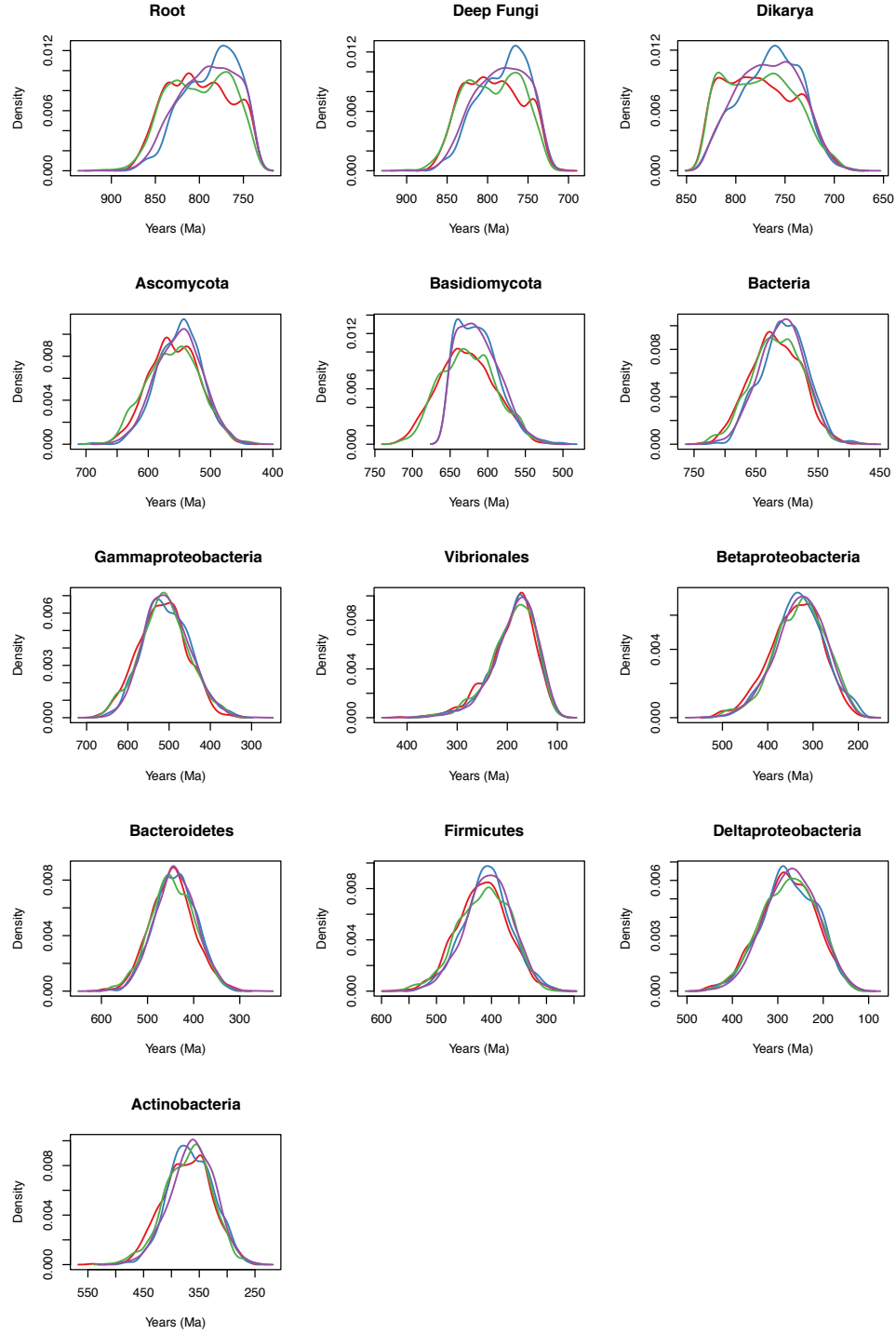


Figure 4-5: Posterior chitinase date distribution across nodes with varied calibration. Illustrates the posterior age distributions across nodes under the four calibration setups (Models 5, 6, 7, 8).

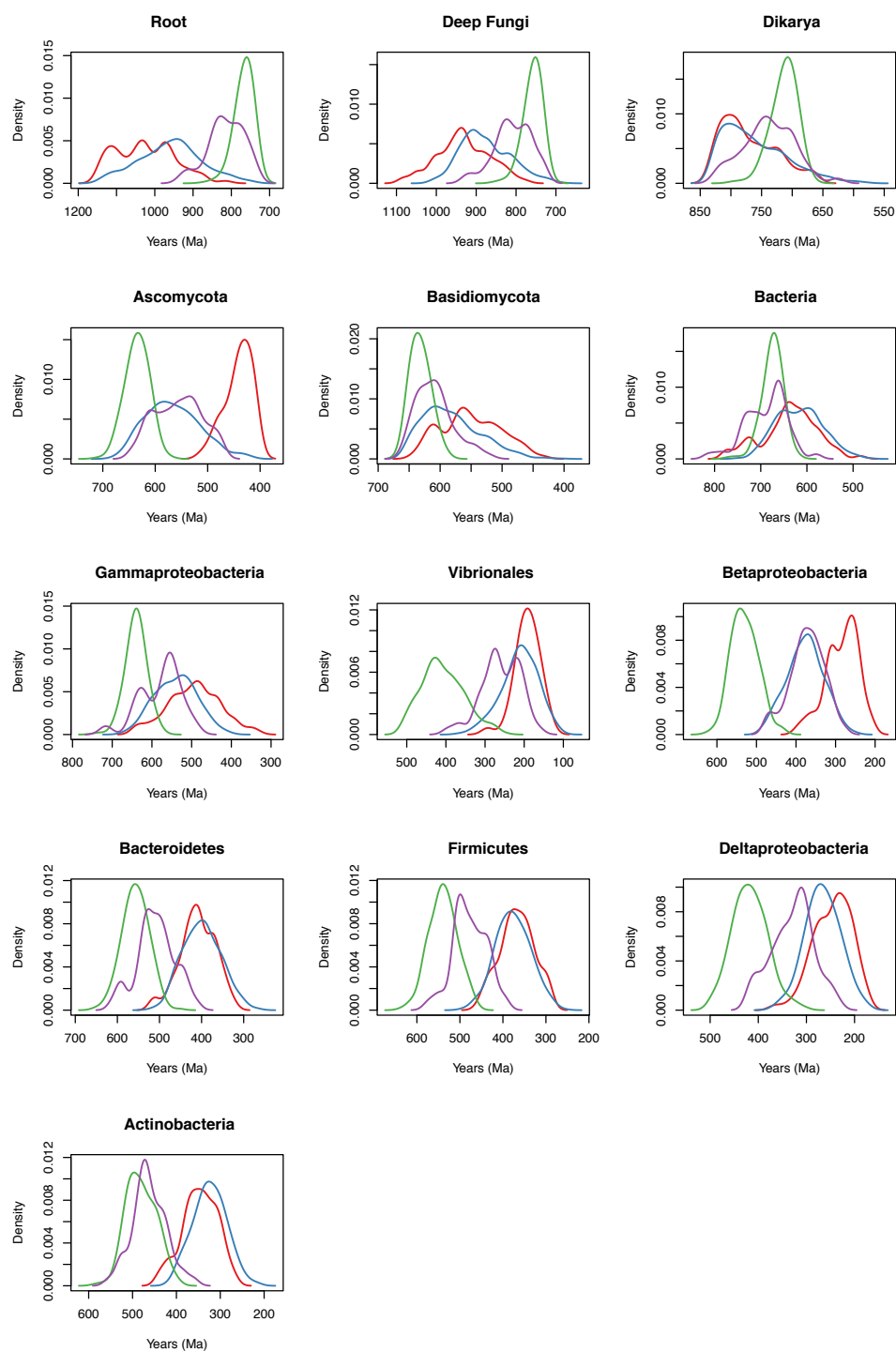


Figure 4-6: Prior date distributions across nodes. Models 2p, 6p, 10p, 14p and Calibration 2.

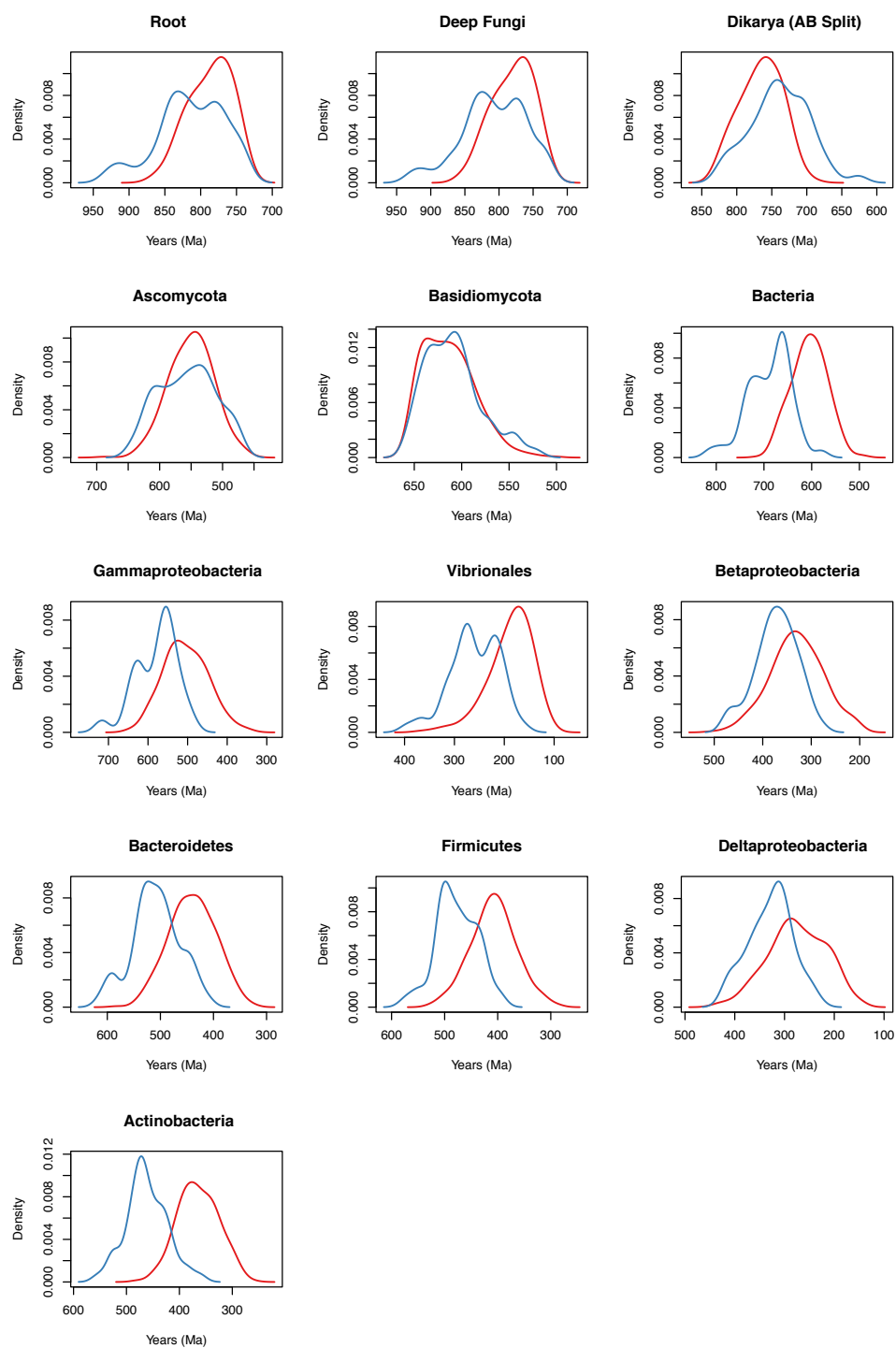


Figure 4-7: Prior vs. posterior chitinase date distributions across nodes.

Table 4-1: Phylobayes model parameters tested in this study. For each model, a \pm indicates the presence or absence of a condition. All models were also tested under the prior (-prior) and sequence data was used to generate posterior probability distributions for all models. BD refers to birth-death. LN stands for lognormal autocorrelated. UGAM stands for uncorrelated gamma multipliers. The AB split refers to the split between Ascomycota and Basidiomycota and fossil refers to the fossil minimum referenced in the Calibration Table 4-2.

Model	Prior		Relaxed Clock Model		Calibrations				
	Uniform	BD	LN	UGAM	Root	AB split	A crown, B crown	Fossil	-prior
1	+	-	+	-	+	+	-	-	-
2	+	-	+	-	+	+	+	-	-
3	+	-	+	-	+	-	-	+	-
4	+	-	+	-	+	-	+	+	-
1p	+	-	+	-	+	+	-	-	+
2p	+	-	+	-	+	+	+	-	+
3p	+	-	+	-	+	-	+	+	+
4p	+	-	+	-	+	-	-	+	+
5	+	-	-	+	+	+	-	-	-
6	+	-	-	+	+	+	+	-	-
7	+	-	-	+	+	-	+	+	-
8	+	-	-	+	+	-	-	+	-
5p	+	-	-	+	+	+	-	-	+
6p	+	-	-	+	+	+	+	-	+
7p	+	-	-	+	+	-	+	+	+
8p	+	-	-	+	+	-	-	+	+
9	-	+	-	+	+	+	-	-	-
10	-	+	-	+	+	+	+	-	-
11	-	+	-	+	+	-	+	+	-
12	-	+	-	+	+	-	-	+	-
9p	-	+	-	+	+	+	-	-	+
10p	-	+	-	+	+	+	+	-	+
11p	-	+	-	+	+	-	+	+	+
12p	-	+	-	+	+	-	-	+	+
13	-	+	+	-	+	+	-	-	-
14	-	+	+	-	+	+	+	-	-
15	-	+	+	-	+	-	+	+	-
16	-	+	+	-	+	-	-	+	-
13p	-	+	+	-	+	+	-	-	+
14p	-	+	+	-	+	+	+	-	+
15p	-	+	+	-	+	-	+	+	+
16p	-	+	+	-	+	-	-	+	+

Table 4-2: Calibrations used in molecular clock models. All calibrations listed in Ma. Models indicates which models (from Table 4-1.) the calibrations were applied to.

Node	Taxon 1	Taxon 2	Calibration	Reference
Root	RozeAll205	Paeniba168	1145-738	Sharpe et al. (2015)
AB Split	SerpLac210	TricVir244	830-518	Floudas et al. (2012)
Ascomycota Crown	RoseNec204	MetaBru136	715-408	Prieto and Wedin (2013)
Basidiomycota Crown	HessVes103	AgarBisp08	655-400	Floudas et al. (2012)
Fossil minima on AB Split	SerpLac210	TricVir244	830-405	Floudas et al. (2012); Berbee and Taylor (2010); Wolfe et al. (2016)

Table 4-3: Posterior divergence time estimates calculated under model assumptions described in Table 4-1 (Model 6). Divergence date ranges are given in Ma.

Node	Node ID	Cal	Mean Age	95% CI
Root	1	1145-738	787	(738-837)
Deep Fungi	3		780	(729-833)
Dikarya (AB Split)	5	830-518	766	(720-828)
Ascomycota	6	715-408	550	(480-618)
Basidiomycota	70	655-400	613	(566-655)
Bacteria	147		605	(537-672)
Gammaproteobacteria	148		505	(393-605)
Vibrionales (A)	165		188	(113-278)
Betaproteobacteria (T)	155		330	(223-442)
Bacteroidetes (A)	174		440	(359-525)
Firmicutes (T)	181		408	(323-498)
Deltaproteobacteria (T)	186		275	(174-389)
Actinobacteria (T)	215		365	(290-435)

Table 4-4: Meta-analysis of posterior date distributions for model parameters. The calibration refers to the node, model number (in parentheses), calibration age (in Ma). The node refers to either Ascomycota (A) or Basidiomycota (B). The prior refers to the tree process prior, either uniform or birth-death. The clock refers to the relaxed clock model, either lognormal autocorrelated (LN) or uncorrelated gamma (ugam). The model refers to the model number as delineated in Table 4-1. The posterior mean age estimates and 95% confidence intervals (CI) are in Ma as calculated in the output of each Phylobayes model. The expected age ranges (Ma) are listed based on literature values as noted in Table 4-2. The Outside Expected (Outside Exp) column lists the percent of the 95% CI outside of the expected range.

Calibration	Node	Prior	Clock	Model	Mean Age	95% CI	Expected	Outside Exp
AB-Split (Cal 1) 830-518	Ascomycota	uniform	LN	1	651	728-580	715-408	8.8
		uniform	ugam	5	559	630-481	715-408	0.0
		BD	LN	13	494	592-391	715-408	8.5
		BD	ugam	9	366	457-267	715-408	74.2
	Basidiomycota	uniform	LN	1	674	749-602	655-400	63.9
		uniform	ugam	5	628	698-559	655-400	30.9
		BD	LN	13	550	650-443	655-400	0.0
		BD	ugam	9	456	572-349	655-400	22.9
AB-Split (Cal 3) 830-405	Ascomycota	uniform	LN	3	656	731-590	715-408	11.3
		uniform	ugam	7	560	637-484	715-408	0.0
		BD	LN	15	510	616-397	715-408	5.0
		BD	ugam	11	366	447-276	715-408	77.2
	Basidiomycota	uniform	LN	3	677	754-607	655-400	67.3
		uniform	ugam	7	626	687-554	655-400	24.1
		BD	LN	15	560	674-430	655-400	7.8
		BD	ugam	11	462	570-369	655-400	15.4

CHAPTER 5

CONCLUDING REMARKS AND FUTURE DIRECTIONS

In summary, the work presented in this thesis involved geochemical and phylogenetic approaches to better understand the evolution and distribution of microbes in the environment. Chapter 2 (and Appendices A, B, and C) focuses on the geochemical aspects of understanding the signals of microbial methanogenesis. I discussed the possible mechanisms that contribute to the kinetic isotope effects observed for the clumped methane isotopologue, $^{13}\text{CH}_3\text{D}$. Values of $\Delta^{13}\text{CH}_3\text{D}$ for pure cultures of methanogens exhibit nonequilibrium isotope effects, unique to enzymatic microbial processes, rather than equilibrium effects exhibited by many thermogenic sources. I discussed how this might be mediated by enzymatic reactions common to all methanogens, rather than specific to each of the three methanogenic pathways. This will be an extremely fruitful area of research in the future, as much is left to be tested in this isotope system.

Chapter 2 represents most of the published data that exist describing $\Delta^{13}\text{CH}_3\text{D}$ for microbial cultures. However, a number of other labs are beginning to culture and measure methane from microbes, so it will be an exciting time to compare results and have more data with which we can understand kinetic isotope effects for $^{13}\text{CH}_3\text{D}$. Looking forward, it will be particularly important to assess other growth conditions, substrates, strains, and isotopologues. Labeling experiments will be incredibly useful for this. Additionally, as mentioned in Chapter 2, methanogens grow at different rates on different substrates. Thus, quantifying and attributing the effects for different processes (e.g. growth rates vs. pathways) will be important. Moreover, it is also of course critical to consider the environmental relevance of these experiments. Appendices B and C touch on possible environmental significance and contributing factors.

Putatively microbial cultures from marine sediments, like methane hydrates sampled in Appendix B, do not show the same "anti-clumped" nonequilibrium isotope values as pure cultures. It is uncertain whether this is due to factors such as methane source processes (e.g. extremely slow growth rates of methanogens in low temperature environments), methane sink processes (e.g. AOM shifting signals in the opposite direction, towards equilibrium as shown in Appendix C), or most likely, some combination of processes and factors not yet described. Co-culturing experiments and experiments representative of environmental conditions will be useful.

Innovative analytical techniques and new technologies have enabled extremely precise measurements of isotopologues in increasingly smaller sample volumes. As this trend continues, more environments and processes will be sampled. Additionally, while the $^{13}\text{CH}_3\text{D}$ equilibrium curve is calibrated at high temperatures by heating processes, lower temperature calibrations are more difficult to make due to the impossibly slow equilibrium exchange reactions at low temperatures. Future work in lower-temperature systems will improve calibrations in this region of the curve ($<100^\circ\text{C}$).

Chapter 3 discusses the possible implications for methylotrophic metabolisms in marine systems and how substrate-specificity may have evolved via HGT. This chapter also highlights a number of hypotheses and future directions to be tested. We readily hypothesize about what marker genes can do, but culturing work will be needed to confirm these hypotheses. It will be important to continue to sample and culture microbes from diverse marine and subsurface environments. Phylogenetically, it will be important to further assess taxon sampling and database annotations, ensuring that results are interpreted correctly. A challenging, but important, next step will be trying to date the emergence of substrate specificity in methylotrophic lineages. This may be aided by ongoing methods developed in Chapter 4.

Chapter 4 describes the discovery of a novel HGT between fungi, bacteria, and within other bacterial lineages. I discuss how the HGT between Fungi and Bacteria may be leveraged to apply fossil calibrations from within fungal lineages to bacterial lineages. As I discuss in this chapter, assumptions about the prior and model selection are incredibly important. This is especially true given that many molecular clock parameters are developed and tested in more recently diverging taxa of different scale, but as we begin the exciting work of applying this to deep microbial divergences.

Results from Chapter 4 suggest that bacterial chitinases originated in fungi and diversified in response to the availability of chitin-producing organisms in the environment. Furthermore, these results suggest that the diversification of terrestrial bacterial chitinases is consistent with that of terrestrial chitin producers (e.g. Arthropods) and a period of possible marine extinction and ecological shifts. While these conclusions are certainly plausible, it will be important to evaluate other chitinases to confirm that this is broadly applicable and assess whether posterior date distributions may be used as secondary calibrations in future work.

As work continues in molecular evolution, we will get increasingly good at interweaving information across the geological and biological sciences to make use of the strongest pieces of evidence. We will also increase the number of tested primary and secondary calibrations to apply to an increasing number of model parameters. I also expect that our databases will not only grow larger, but that we will get better at annotation and curation, such that we can make use of the growing number of sequenced genes and genomes. I look forward to improved usability and interoperability between phylogenetics software and file formats which will increase research output in the field. Finally, I am excited to see how the field of geobiology continues to build upon emerging molecular evolution methods, as early microbial evolution remains an exciting frontier in biology. Geochemical, biological, and computational tools together will help to illuminate some of the most temporally and geographically remote environments.

APPENDIX A

NONEQUILIBRIUM CLUMPED ISOTOPE SIGNALS IN MICROBIAL METHANE

Wang DT, **Gruen DS**, Lollar BS, Hinrichs K-U, Stewart LC, Holden JF, Hristov AN, Pohlman JW, Morrill PL, Könneke M, Delwiche KB, Reeves EP, Sutcliffe CN, Ritter DJ, Seewald JS, McIntosh JC, Hemond HF, Kubo MD, Cardace D, Hoehler TM, Ono S (2015). Nonequilibrium clumped isotope signals in microbial methane. *Science* 348:428-431.

D.T.W. and S.O. developed the methods, analyzed data, and performed modeling. D.T.W. and D.S.G. performed isotopic analyses. D.S.G., L.C.S., J.F.H., M.K., K.U.H., and S.O. designed and/or conducted microbiological experiments. D.T.W., D.S.G., B.S.L., P.L.M., K.B.D., A.N.H., C.N.S., M.D.K., D.J.R., J.C.M., D.C., and S.O. designed and/or executed the field sampling campaigns. D.T.W. and S.O. wrote the manuscript with input from all authors.

Methane is a key component in the global carbon cycle with a wide range of anthropogenic and natural sources. Although isotopic compositions of methane have traditionally aided source identification, the abundance of its multiply-substituted "clumped" isotopologues, *e.g.*, $^{13}\text{CH}_3\text{D}$, has recently emerged as a proxy for determining methane-formation temperatures; however, the impact of biological processes on methane's clumped isotopologue signature is poorly constrained. We show that methanogenesis proceeding at relatively high rates in cattle, surface environments, and laboratory cultures exerts kinetic control on $^{13}\text{CH}_3\text{D}$ abundances and results in anomalously elevated formation temperature estimates. We demonstrate quantitatively that H_2 availability accounts for this effect. Clumped methane thermometry can therefore provide constraints on the generation of methane in diverse settings, including continental serpentinization sites and ancient, deep groundwaters.

Carbon ($^{13}\text{C}/^{12}\text{C}$) and hydrogen (D/H) isotope ratios of methane are widely applied for distinguishing microbial from thermogenic methane in the environment (Baldassare et al., 2014; Flores et al., 2008; Pohlman et al., 2009; Sherwood Lollar et al., 2008; Sherwood Lollar et al., 2002; Welhan and Lupton, 1987; Whiticar, 1990) as well as for apportioning pathways of microbial methane production (Burke Jr et al., 1988; McCalley et al., 2014; Whiticar et al., 1986). This bulk isotope approach, however, is largely based on empirical observations, and different origins of methane often yield overlapping characteristic isotope signals (Pohlman et al., 2009; Whiticar, 1990; Etiope and Sherwood Lollar, 2013; Schoell, 1988; Whiticar, 1999). Beyond conventional bulk isotope ratios, it has become possible to precisely measure the abundance of multiply-substituted "clumped" isotopologues (*e.g.*, $^{13}\text{CH}_3\text{D}$) (Ono et al., 2014; Stolper et al., 2014). In particular, abundance of clumped isotopes promises to yield information about the temperature at which C_H bonds were formed or last equilibrated (Ono et al., 2014, ; Fig. S1). Indeed, formation temperatures of both thermogenic and microbial methane in natural gas reservoirs can be estimated on the basis of clumped isotopologues (Stolper et al., 2014). The mechanisms by which isotopologues attain distributions consistent with thermodynamic equilibrium, however, remain unclear because bulk methane isotopes ($\delta^{13}\text{C}$ and δD) often reflect kinetic isotope fractionations (Whiticar, 1999; Valentine et al., 2004), and H-isotope exchange between methane and water is sluggish (Reeves et al., 2012).

To test if clumped methane thermometry can be widely applied for methane sources beyond natural gas reservoirs, we examined methane samples from diverse systems, including lakes, wetlands, cow rumen, laboratory cultures of methanogenic microbes, and geological settings that may support abiogenic methane production as well as thermogenic and microbial sources, including continental serpentinization sites and deep fracture fluids. We measured the relative abundances of four methane isotopologues ($^{12}\text{CH}_4$, $^{13}\text{CH}_4$, $^{12}\text{CH}_3\text{D}$ and $^{13}\text{CH}_3\text{D}$) using a recently-developed tunable laser spectroscopy technique (Ono et al.,

2014, and see Supplemental Materials).

Our measurements for dominantly-thermogenic gases from the Marcellus and Utica Shales (Baldassare et al., 2014; Burruss and Laughrey, 2010) yielded $\Delta^{13}\text{CH}_3\text{D}$ -based temperatures of $147^{+25}_{-22}^\circ\text{C}$ and $160^{+29}_{-25}^\circ\text{C}$, respectively. The clumped-isotope temperature for the Marcellus Shale sample is comparable to, although slightly lower than, estimates by Stolper et al. (2014) of $179\text{--}207^\circ\text{C}$ (Fig. A-1). In addition, microbial methane in pore waters and gas hydrates from northern Cascadia margin sediments (Pohlman et al., 2009), and from wells producing from coal seams in the Powder River Basin (Flores et al., 2008; Bates et al., 2011) yielded $\Delta^{13}\text{CH}_3\text{D}$ temperatures of $12\text{--}42^\circ\text{C}$ and $35\text{--}52^\circ\text{C}$, respectively. These are consistent with their expected low formation temperatures. Furthermore, thermogenic methane sampled from a hydrothermal vent in the Guaymas Basin, Gulf of California (Welhan and Lupton, 1987), yielded $\Delta^{13}\text{CH}_3\text{D}$ temperature of $326^{+170}_{-95}^\circ\text{C}$, within error of the measured vent temperature (299°C Reeves et al., 2014). Therefore, our data provide independent support of the hypothesis that $^{13}\text{CH}_3\text{D}$ abundance reflects the temperature at which methane is generated in these sedimentary basins (Stolper et al., 2014).

In contrast, we found that methane sampled from lakes, a swamp, and the rumen of a cow carry $^{13}\text{CH}_3\text{D}$ signals that correspond to anomalously high $\Delta^{13}\text{CH}_3\text{D}$ temperatures ($139\text{--}775^\circ\text{C}$, Fig. A-1A), *i.e.*, well above the environmental temperatures ($<40^\circ\text{C}$). Such signals are clearly not controlled by equilibrium. Notably, a positive correlation between $\Delta^{13}\text{CH}_3\text{D}$ and the extent of D/H fractionation between methane and environmental water [$\epsilon_{\text{methane/water}}$ (23); Fig. A-2] suggests a strong link between isotopologue (*i.e.*, $^{13}\text{CH}_3\text{D}$) and isotope (D/H) disequilibria. In contrast, the above mentioned methane samples from sedimentary basins appear to have attained hydrogen-isotope equilibrium with associated waters at or near the temperatures indicated by the $\Delta^{13}\text{CH}_3\text{D}$ data (Fig. A-2).

To confirm these observations from the natural environment, we demonstrated that strong disequilibrium $^{13}\text{CH}_3\text{D}$ signals are also produced by cultures of methanogenic archaea in the laboratory (Fig. A-3). Thermophilic methanogens cultured at 40 to 85°C produced methane with $\Delta^{13}\text{CH}_3\text{D}$ values from $+0.5$ to $+2.3\text{‰}$ (corresponding to $\Delta^{13}\text{CH}_3\text{D}$ temperatures of $216\text{--}620^\circ\text{C}$), and mesophilic methanogens cultured at ambient temperature produced methane with conspicuously "anti-clumped" signatures (*i.e.*, values of $\Delta^{13}\text{CH}_3\text{D} < 0\text{‰}$, for which no apparent temperature can be expressed) as low as -1.3‰ (Fig. A-3). Methane from cultures is also characterized by large kinetic D/H fractionation with respect to water (Valentine et al., 2004; Balabane et al., 1987). Because laboratory cultures are grown under optimal conditions (high- H_2 and high- CO_2), these anti-clumped $\Delta^{13}\text{CH}_3\text{D}$ and low $\epsilon_{\text{methane/water}}$ values are primarily expressions of kinetic isotope effects. Consequently, the distribution of samples with $\Delta^{13}\text{CH}_3\text{D}$ and $\epsilon_{\text{methane/water}}$ values in Fig. A-2 can be explained by microbial methanogenesis operating on a spectrum between fully kinetic (low $\Delta^{13}\text{CH}_3\text{D}$ and low $\epsilon_{\text{methane/water}}$) and equilibrium (high $\Delta^{13}\text{CH}_3\text{D}$ and high $\epsilon_{\text{methane/water}}$)

end-members.

We constructed a mathematical framework to describe the controls on the correlation of $\Delta^{13}\text{CH}_3\text{D}$ and $\epsilon_{\text{methane}/\text{water}}$ signals from hydrogenotrophic methanogenesis. The model largely follows those developed for microbial sulfate reduction (Rees, 1973; Wing and Halevy, 2014) and predicts the isotopologue compositions of product methane as a result of a series of enzymatic reactions [Fig. S4; available online]. Using isotope fractionation factors estimated from theory, experiments and observations as input parameters [Table S3; available online], our model reproduces the observed correlation between $\Delta^{13}\text{CH}_3\text{D}$ and $\epsilon_{\text{methane}/\text{water}}$ of natural samples (Fig. A-2). The isotopologue compositions of product methane reflect the degree of metabolic reversibility. Fully reversible reactions yield equilibrium end-members (Holler et al., 2011), while irreversible reactions result in kinetic (disequilibrium) end-member signals. In this model, the reversibility is linked to available free energy (Wing and Halevy, 2014; Holler et al., 2011), in this case expressed as H_2 concentration ($[\text{H}_2]$). The model can explain the relationship among $[\text{H}_2]$, $\epsilon_{\text{methane}/\text{water}}$ (Burke, 1993) and $\Delta^{13}\text{CH}_3\text{D}$ via Michaelis-Menten kinetics, and predicts the observed patterns in diverse settings ranging from marine sediments (low $[\text{H}_2]$, high $\Delta^{13}\text{CH}_3\text{D}$ and $\epsilon_{\text{methane}/\text{water}}$) to bovine rumen (high $[\text{H}_2]$, low $\Delta^{13}\text{CH}_3\text{D}$ and $\epsilon_{\text{methane}/\text{water}}$) (Fig. A-4). We note that mixing of methane sources with different $\delta^{13}\text{C}$ and δD values or oxidation of methane could also alter the relationships over the primary signal of microbial methanogenesis [See supplement online]. Likewise, inheritance of clumping signals from precursor organic substrates (*e.g.*, via acetoclastic or methylotrophic methanogenesis), cannot be entirely ruled out and await experimental validation.

We showed above that the combination of $\Delta^{13}\text{CH}_3\text{D}$ and $\epsilon_{\text{methane}/\text{water}}$ values provides mechanistic constraints on whether methane was formed under kinetic vs. near-equilibrium conditions. Next, we used this framework to place constraints on the origins of methane at two sites of present-day serpentinization in Phanerozoic ophiolites (The Cedars and Coast Range Ophiolite Microbial Observatory, CROMO Morrill et al., 2013; Cardace et al., 2013) in northern California, and in deep (> 2 km below surface) fracture fluids with billion year-residence times in the Kidd Creek mine, Canada (Sherwood Lollar et al., 2002; Holland et al., 2013).

Methane-rich gases in groundwater springs associated with serpentinization at The Cedars yielded anti-clumped $\Delta^{13}\text{CH}_3\text{D}$ signals ($\sim 3\text{‰}$) with low $\epsilon_{\text{methane}/\text{water}}$ values (Figs. A-1A and A-2). The data plot along the microbial (kinetic) trend defined in Fig. A-2, supporting a previous hypothesis that methane at The Cedars is being produced by active microbial methanogenesis (Morrill et al., 2013). The exceptionally high H_2 concentration (up to 50% by volume in bubbles) and low Eh (ca. ~ 600 mV) at The Cedars indicate the massive excess of electron donor. This, along with severe inorganic carbon limitation (due to high pH (> 11) and precipitation of carbonate minerals Morrill et al., 2013), drives the forma-

tion of methane carrying strong kinetic imprints, consistent with the observed anti-clumped $\Delta^{13}\text{CH}_3\text{D}$ signals (Fig. A-4).

Despite the similarity in geologic setting, methane associated with serpentinization at CROMO (Cardace et al., 2013) revealed very different $\Delta^{13}\text{CH}_3\text{D}$ values, which correspond to low apparent temperatures (42-76°C) and plot close to the equilibrium line (Fig. A-2). While the conventional $\delta^{13}\text{C}$ and δD values of methane from CROMO are nearly identical to those of the Utica Shale sample (Fig. A-1B), methane at CROMO carries much higher $\Delta^{13}\text{CH}_3\text{D}$ values (Fig. A-1A). The origin of methane at the CROMO site remains unresolved (Cardace et al., 2013), but the comparably high $\Delta^{13}\text{CH}_3\text{D}$ values at CROMO suggest methane here could be sourced from a mixture of thermogenic and microbial methane. Alternatively, lower H_2 availability at CROMO, compared to The Cedars (table S4; available online), may support microbial methanogenesis under near-equilibrium conditions (Fig. A-4). Regardless, the different isotopologue signatures in methane from CROMO vs. The Cedars demonstrate that distinct processes contribute to methane formation in these two serpentinization systems.

Deep, ancient fracture fluids in the Kidd Creek mine in the Canadian Shield (Holland et al., 2013) contain copious quantities of both dissolved methane and hydrogen (Sherwood Lollar et al., 2002). The Kidd Creek methane occupies a distinct region in the $\Delta^{13}\text{CH}_3\text{D}$ vs. $\epsilon_{\text{methane/water}}$ diagram (Fig. A-2), due to strong D/H disequilibria between methane and water (Sherwood Lollar et al., 2008) and low $\Delta^{13}\text{CH}_3\text{D}$ temperature signals of 56-90 °C that are consistent with other temperature estimates for these groundwaters (Sherwood Lollar et al., 2008). Although the specific mechanisms by which the proposed abiogenic hydrocarbons at Kidd Creek are generated remain under investigation (Sherwood Lollar et al., 2002; Sherwood Lollar et al., 2014), the distinct isotopologue signals provide further support for the hypothesis that methane here is neither microbial nor thermogenic.

Our results demonstrate that measurements of $^{13}\text{CH}_3\text{D}$ provide information beyond the simple formation temperature of methane. Combination of methane/water hydrogen-isotope fractionation and $^{13}\text{CH}_3\text{D}$ abundance enables the differentiation of methane that has been formed at extremely low rates in the subsurface (Pohlman et al., 2009; Bates et al., 2011; Holler et al., 2011) from methane formed in cattle and surface environments in which methanogenesis proceeds at comparatively high rates (Johnson and Johnson, 1995; Varadharajan and Hemond, 2012).

Supplementary methods, text, figures, tables and references associated with this article can be found online here.

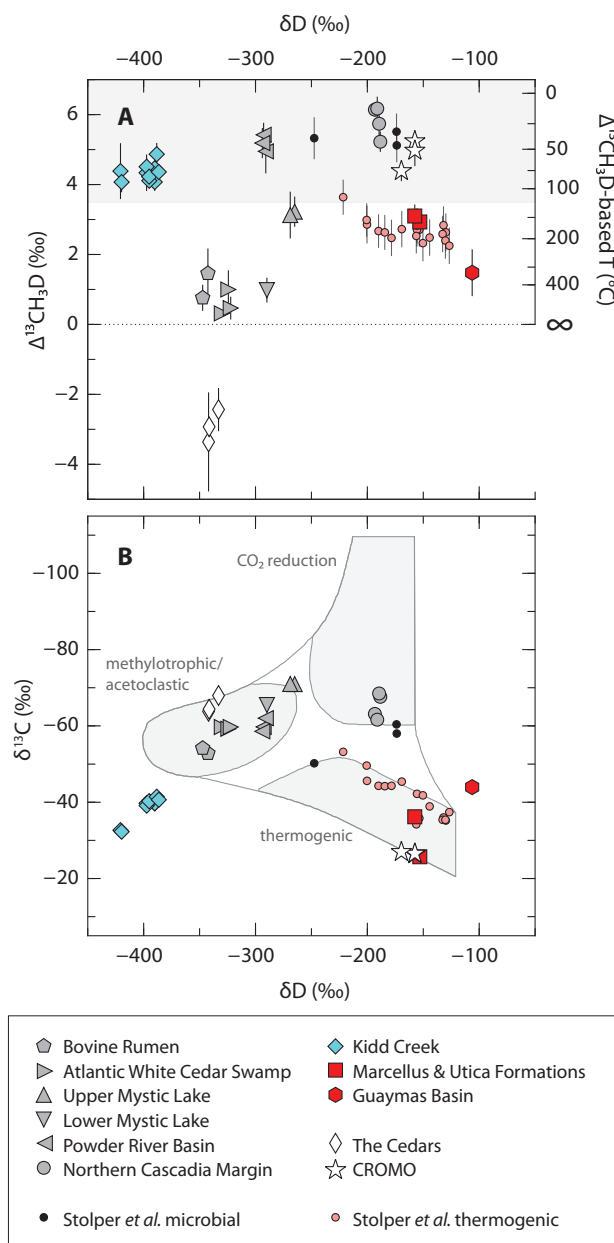


Figure A-1: Isotopologue compositions of methane samples. (A) $\Delta^{13}\text{CH}_3\text{D}$ plotted against δD . The $\Delta^{13}\text{CH}_3\text{D}$ temperature scale corresponds to calibration in fig. S1. Error bars are 95% confidence intervals (table S1; available online). Data from (Stolper *et al.*, 2014) were scaled to their corresponding $\Delta^{13}\text{CH}_3\text{D}$ values. The shaded area represents the temperature range within which microbial life has been demonstrated to date (Takai *et al.*, 2008). The hatched line represents $\Delta^{13}\text{CH}_3\text{D} = 0$ ‰ ($T \rightarrow \infty$); data plotting below this line cannot yield corresponding apparent temperatures. (B) $\delta^{13}\text{C}$ plotted against δD , showing characteristic fields for different methane sources from (Whiticar, 1999).

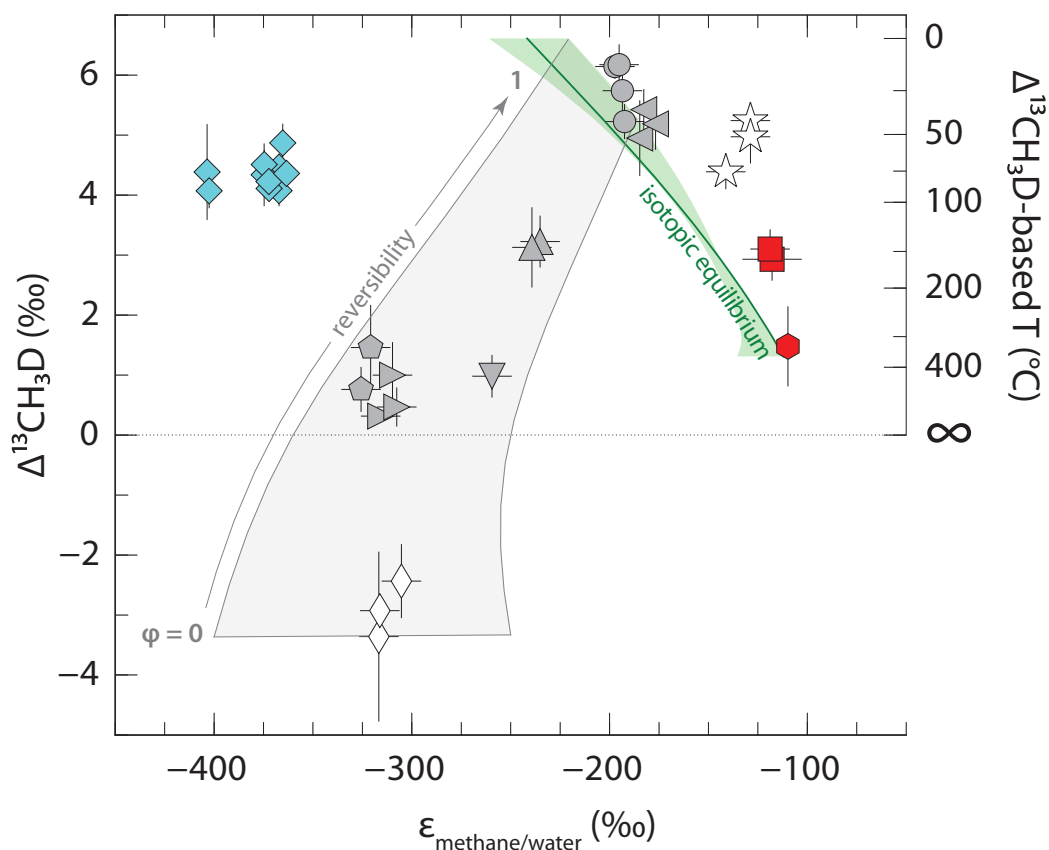


Figure A-2: Extent of clumped- and hydrogen-isotopic disequilibria in methane. Symbols and vertical error bars are the same as those in Fig. A-1. Horizontal error bars represent uncertainties on estimates of $\epsilon_{\text{methane/water}}$ [table S4]. The solid green curve represents isotopic equilibrium, with the $\epsilon_{\text{methane/water}}$ calibration given by (Horibe and Craig, 1995). Green shading represents ranges of $\epsilon_{\text{methane/water}}$ calibrations from published reports (fig. S3). Gray shading represents model predictions from this study, for microbial methane formed between 0 and 40 °C. Metabolic reversibility (Φ) increases from bottom ($\Phi = 0$, fully-kinetic) to top ($\Phi \rightarrow 1$, equilibrium) within this field (see Supplement).

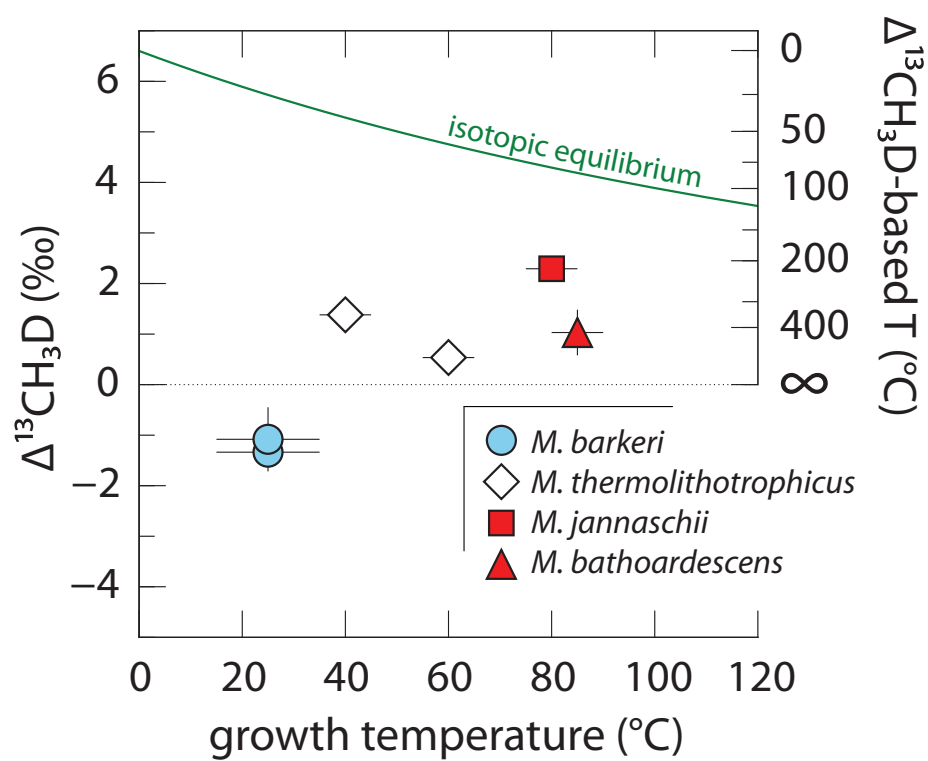


Figure A-3: $\Delta^{13}\text{CH}_3\text{D}$ values of methane produced by hydrogenotrophic methanogens in batch cultures reflect kinetic effects. Data and error bars are from table S2. The green line represents clumped isotopologue equilibrium (*i.e.*, samples for which $\Delta^{13}\text{CH}_3\text{D}$ temperature is equal to growth temperature; fig. S1).

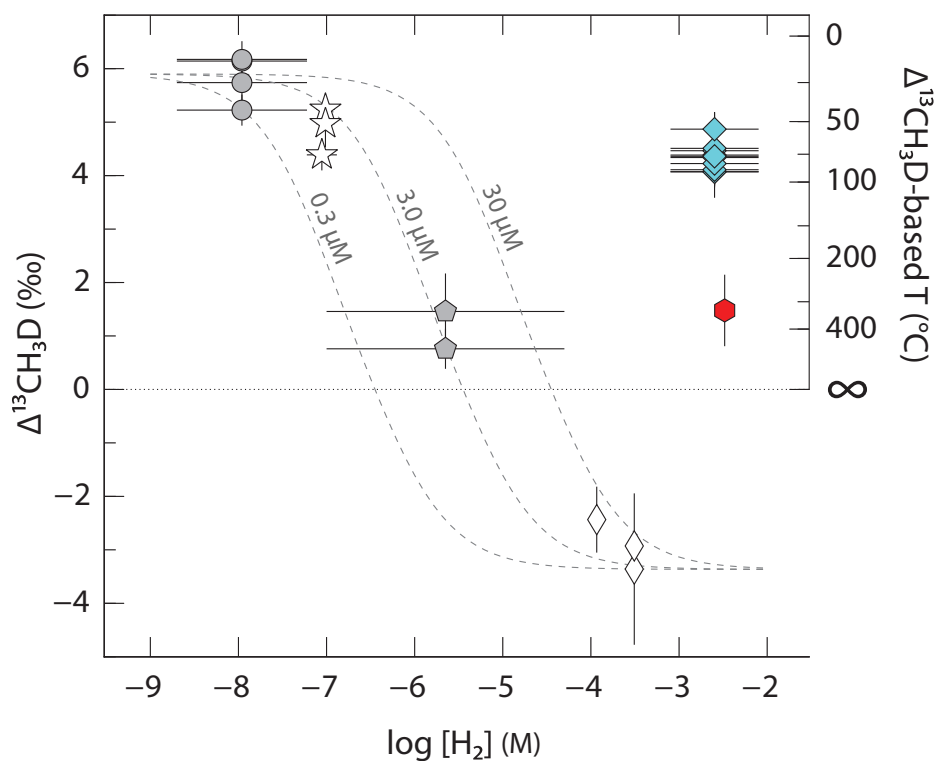


Figure A-4: Relationships between $\Delta^{13}\text{CH}_3\text{D}$ and H_2 concentration for microbial methane. Symbols and vertical error bars are the same as in Fig. A-1. The H_2 data are from table S4; when a range of $[\text{H}_2]$ values is given, points are plotted at the geometric mean of the maximum and minimum values. Dashed lines represent model predictions for microbial methane produced at 20 °C, calculated using KM's of 0.3, 3.0, and 30 $\mu\text{M H}_2$. Data for samples of non-dominantly-microbial methane from Guaymas Basin and Kidd Creek are plotted for comparison.

APPENDIX B

METHANE CLUMPED ISOTOPOLOGUE ($^{13}\text{CH}_3\text{D}$) MEASUREMENTS FOR DETERMINING THE MIXING OF MICROBIAL AND THERMOGENIC METHANE IN OCEANIC GAS HYDRATES

Motivation and Outlook

Methane hydrates make up one of the largest reservoirs of methane on Earth (Kvenvolden, 1993). Most methane hydrates are considered to be microbial (Kvenvolden, 1993), archives of deep biosphere microbial methane (Inagaki et al., 2006; Valentine et al., 2004). Methanogens are the only organisms capable of producing methane as a catabolic end product (Valentine, 2011). In anoxic ocean sediments, organic matter is decomposed and microbes reduce the resulting CO_2 via methanogens to form methane (Thauer et al., 2008). This methane is trapped by the water that crystallizes to ice, forming sedimentary hydrate outcrops within the gas hydrate stability zone (Pohlman et al., 2009). Because this biosphere is difficult to sample and study, very little is known about how, where, or when microbes generate this methane. Moreover, these systems can be quite dynamic, with mixing of multiple methane sources (Milkov et al., 2005) and environmental restructuring (Sultan et al., 2014). Complex architectures of hydrate-bearing sediments can form, dissolve, and evolve over geologically-relevant timescales (Sassen et al., 2001; Suess et al., 1999). It is likely that these systems do preserve microbial methane, but rather than being stagnant archives, they are more dynamic and complex than was once thought. Conventional carbon and hydrogen stable isotopes are often used to determine the source of methane as a proxy in remote environments (Heuer et al., 2009). However, these measurements produce often overlapping signals, confounded by formation processes and source values (Chp. 2 Gruen et al., 2018). Novel measurements of methane isotopologues may aid in determining the formation process (*e.g.*, microbial vs. thermogenic) for methane in these environments, particularly due to differences in how these molecules form and mix (Chp. 2, App. A; Gruen et al., 2018; Stolper et al., 2013, 2014, 2015; Wang et al., 2015). It is possible that these measurements will indicate that some environments may have a greater contribution from microbial sources than was once thought. The data in this appendix represent data that was sampled and measured but not incorporated into the main thesis chapters.

Acknowledgments

My sincerest thanks to Gerhard Bohrmann and Thomas Pape at the MARUM, Bremen, Germany for inspiring me to work on these samples, teaching me so much about oceanic gas hydrates, and sharing your precious hydrates with the Ono lab. Thank you to John Pohlman, USGS, Verena Heuer, MARUM, Andrew Whitehill, and Shuhei Ono for productive discussions on hydrates and source mixing. Thank you to Jen Karolewski, MIT-WHOI Joint Program PhD student, for helping with isotope analyses.

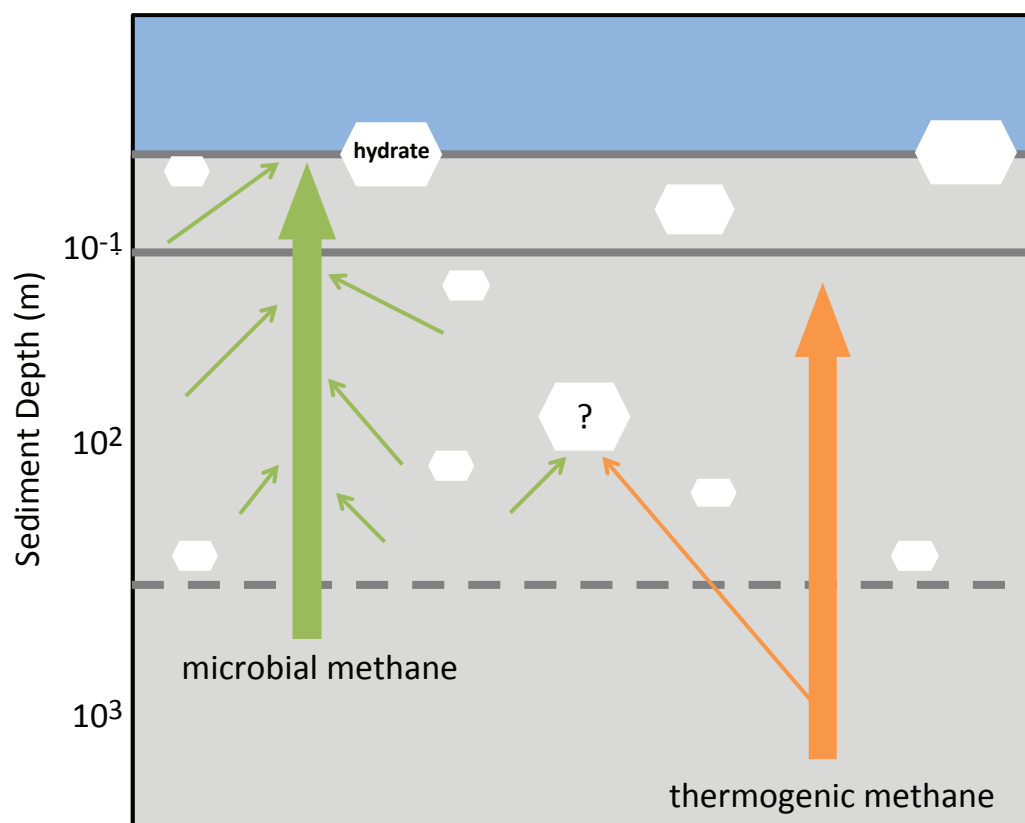


Figure B-1: Schematic of methane hydrate formation modified from (Pohlman et al., 2009).

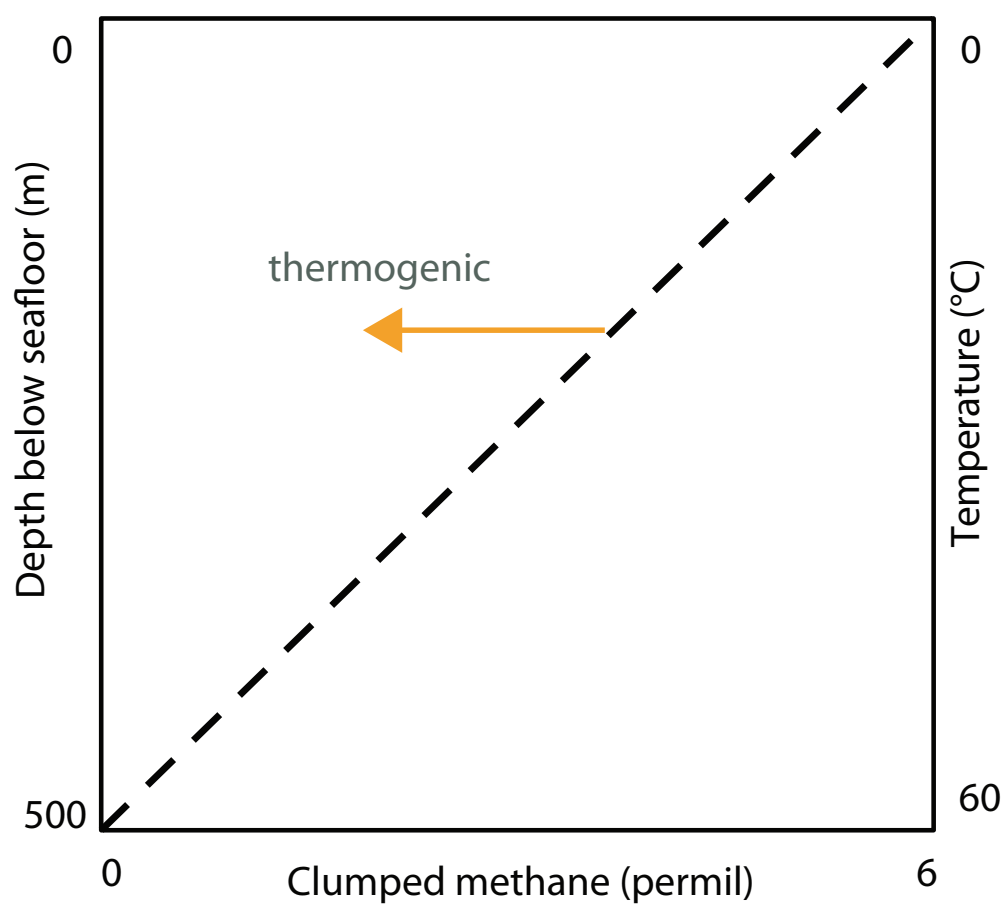


Figure B-2: Hypothesized relationship between sediment depth and clumped methane.

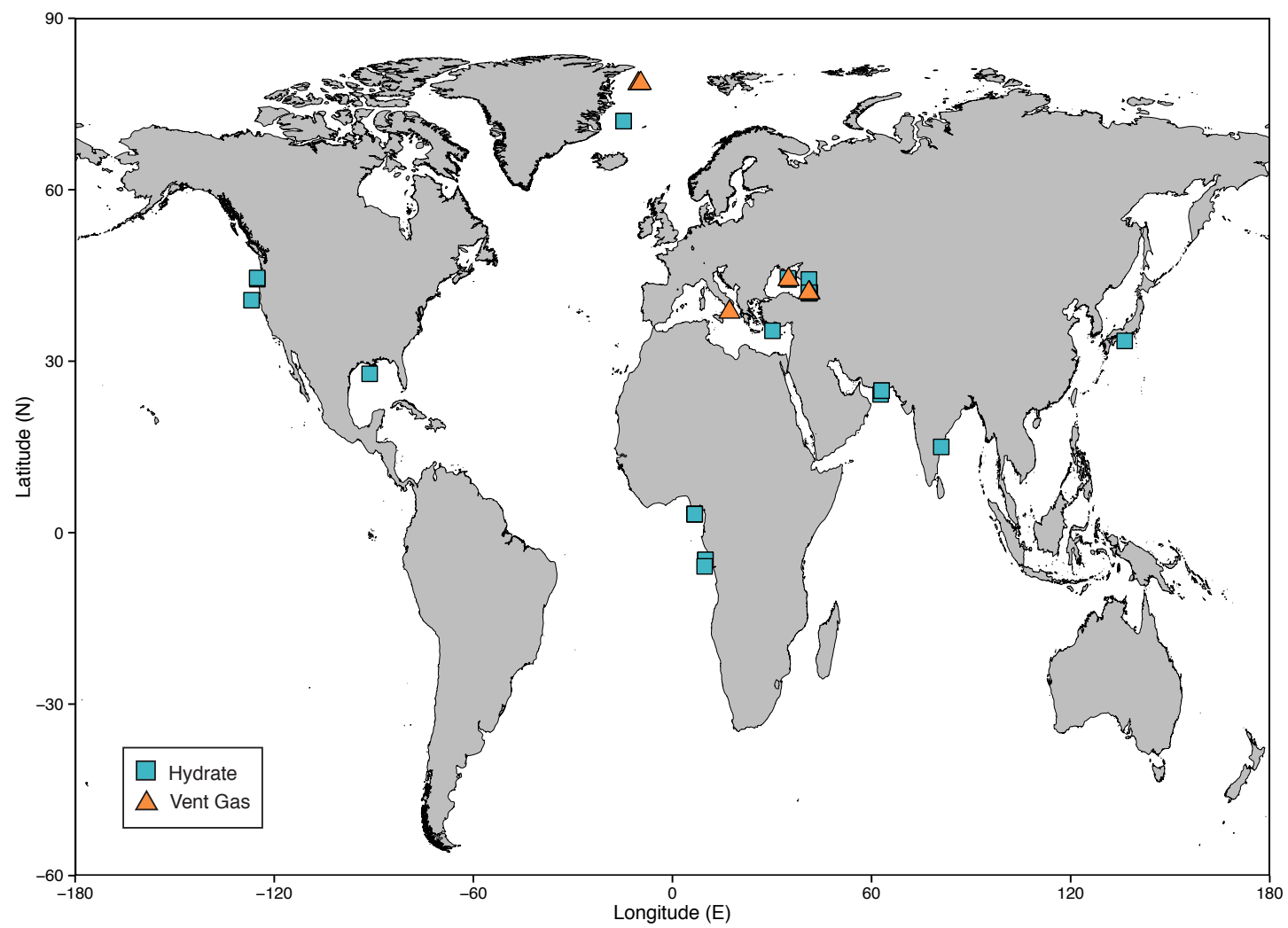


Figure B-3: Sample site locations.

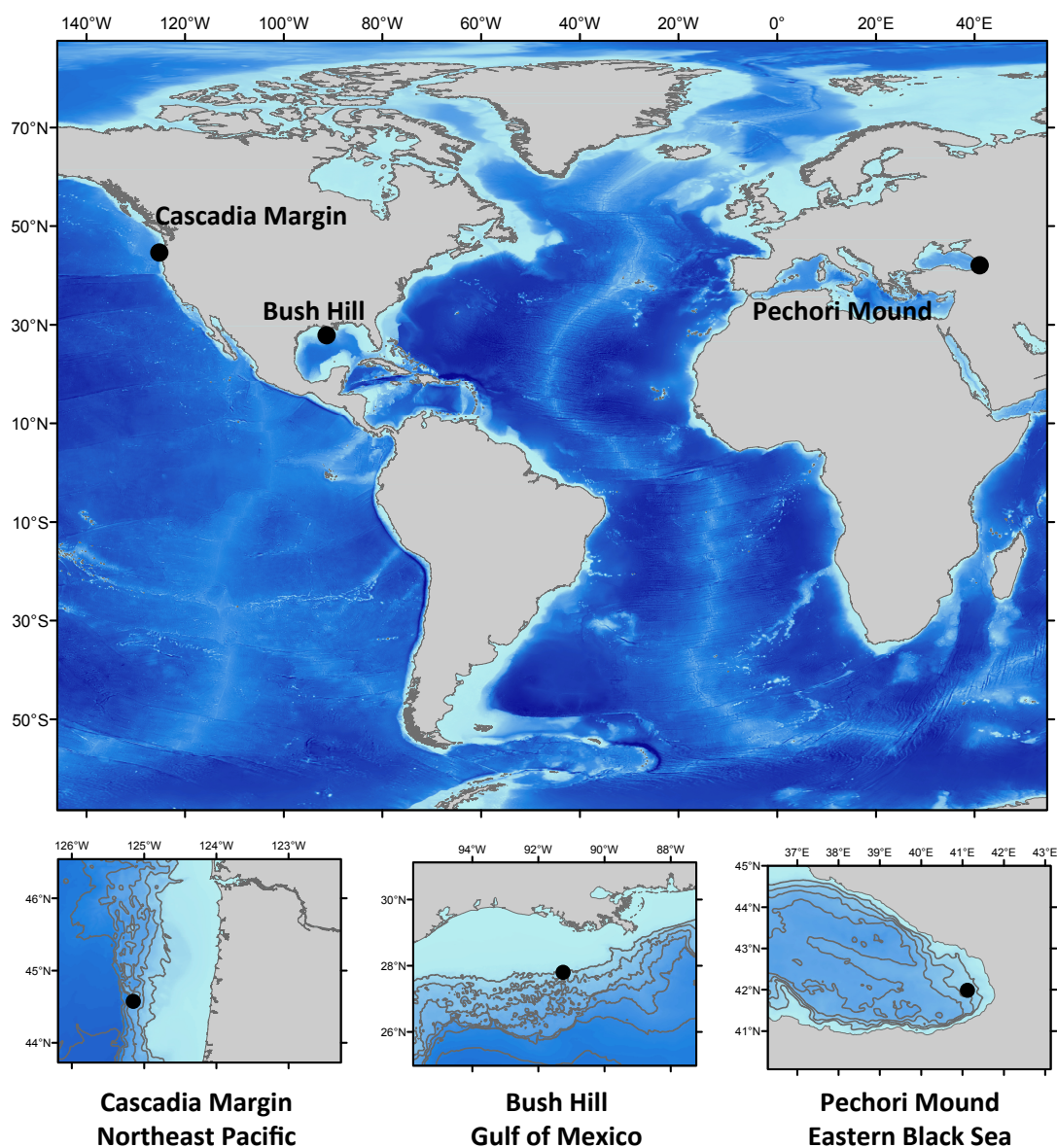


Figure B-4: Three sample sites highlighted. Cascadia Margin is putatively microbial, Bush Hill is putatively thermogenic, and Pechori Mound is putatively microbial.

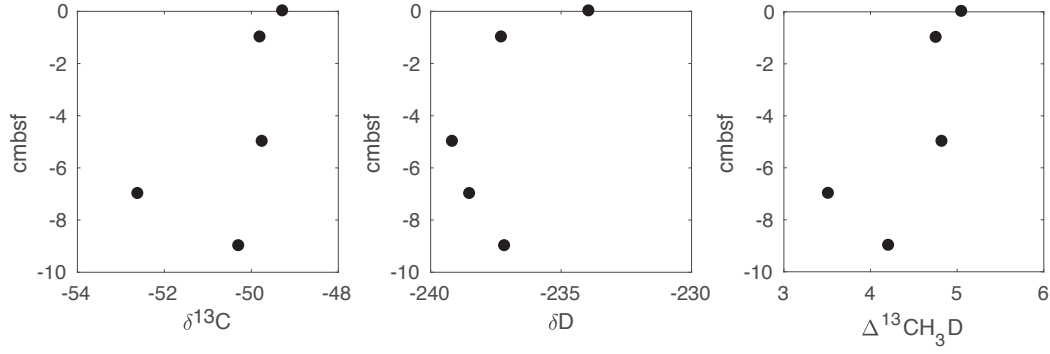


Figure B-5: Depth profile of $\delta^{13}\text{C}$, δD , and $\Delta^{13}\text{CH}_3\text{D}$ at the Pechori Mound site, Eastern Black Sea.

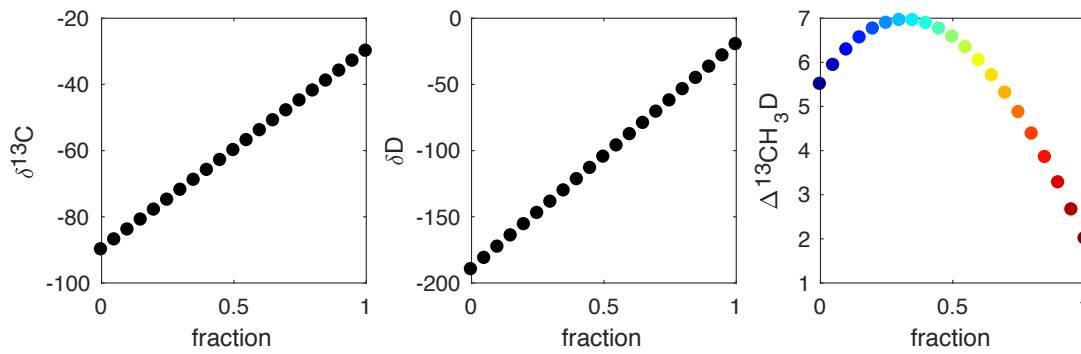


Figure B-6: Mixing Model illustrating the relationship between the mixing of two sources of $\delta^{13}\text{C}$, δD , and highlighting the nonlinearity of mixing $\Delta^{13}\text{CH}_3\text{D}$ of two sources.

Table B-1: Sample descriptions and associated references.

GeoB	Site	Tool	Region	Gas type	C1/C2	$\delta^{13}\text{C-CH}_4$ [‰V-PDB]	Literature
Black Sea							
15260	Batumi Seep	Gravity Corer	Eastern Black Sea	hydrate-bound gas	4,178	-53.8	Pape et al. (2011); Reitz et al. (2011)
11907	Batumi Seep	Gas Bubble Sampler	Eastern Black Sea	vent gas	5,383	-53.2	Pape et al. (2011); Reitz et al. (2011)
11921-1	Batumi Seep	Gas Bubble Sampler	Eastern Black Sea	vent gas	4,631	-51.5	
11971	Colkhети Seep	Gravity Corer	Eastern Black Sea	hydrate-bound gas	32	-49.5	Reitz et al. (2011); Körber et al. (2014)
11902-1	Colkhети Seep	Gas Bubble Sampler	Eastern Black Sea	vent gas	190	-48.1	Reitz et al. (2011); Körber et al. (2014)
11938	Iberia Mound	Gravity Corer	Eastern Black Sea	hydrate-bound gas	2,090	-53.1	Reitz et al. (2011); Körber et al. (2014)
15268-1	Ordu ridge patch	Gravity corer	Eastern Black Sea	hydrate-bound gas	3,131	-70.8	
15503-1	Ordu ridge patch	Gravity corer	Eastern Black Sea	hydrate-bound gas	2,816	-70.4	
15505	Ordu ridge patch	Gravity corer	Eastern Black Sea	hydrate-bound gas	2,335	-72.6	
15507	Ordu ridge patch	Gravity corer	Eastern Black Sea	hydrate-bound gas	3,258	-71.3	
15227-3	Pechori Mound	MeBo	Eastern Black Sea	hydrate-bound gas	5	-48.8	Reitz et al. (2011); Körber et al. (2014)
15244-2	Poti Seep	Gravity corer	Eastern Black Sea	hydrate-bound gas	4,153	-50.3	
11913	Vodyanitskii MV	Gravity corer	Northern Black Sea	hydrate-bound gas	2,018	-59.2	Sahling et al. (2009)
15525-1	Hegoland MV	Gravity Corer	Northern Black Sea	hydrate-bound gas	3,054	-63.6	
14339-3	Hegoland MV	Gas Bubble Sampler	Northern Black Sea	vent gas	2,257	-62.8	
15518	Kerch Flare	Gravity Corer	Northern Black Sea	hydrate-bound gas	2,498	-66.9	Römer et al. (2012)
Continental margin W of Africa							
16001-1	Pockmark_A	Gravity Corer	Gulf of Guinea, S of Nigeria	hydrate-bound gas	6,556	-51.1	Sultan et al. (2014); Wei et al. (2015)
16022-1	Pockmark_A	Gravity Corer	Gulf of Guinea, S of Nigeria	hydrate-bound gas	8,443	-51	Sultan et al. (2014); Wei et al. (2015)
16005-2	Pockmark_C1	Gravity Corer	Gulf of Guinea, S of Nigeria	hydrate-bound gas	4,060	-53	Sultan et al. (2014); Wei et al. (2015)
16016-1	Pockmark_C1	Gravity Corer	Gulf of Guinea, S of Nigeria	hydrate-bound gas	6,467	-51.3	Sultan et al. (2014); Wei et al. (2015)
13114-3	Hydrate Hole	Gravity Corer	Northern Congo Fan	hydrate-bound gas	1,988	-70.6	Sahling et al. (2008)
13115-1	Baboon Hole	Gravity Corer	Northern Congo Fan	hydrate-bound gas	1,638	-69.3	Sahling et al. (2008)
13118-1	Worm Hole	Gravity Corer	Northern Congo Fan	hydrate-bound gas	1,419	-70.7	Sahling et al. (2008)
13120-4	Deep Hole	Gravity Corer	Northern Congo Fan	hydrate-bound gas	1,629	-70.8	Sahling et al. (2008)
13121-1	Regab pockmark	Gravity Corer	Northern Congo Fan	hydrate-bound gas	1,599	-68.4	Sahling et al. (2008)
SW of Japan							
16716-2	MV10	Gravity Corer	Kumano Basin, S of Japan	hydrate-bound gas	65	-37.7	Pape et al. (2014)
16736-2	MV4	Gravity Corer	Kumano Basin, S of Japan	hydrate-bound gas	59	-39.3	Pape et al. (2014)
16746	MV5	Gravity Corer	Kumano Basin, S of Japan	hydrate-bound gas	1,199	-38.3	Pape et al. (2014)
16772	MV2	Gravity Corer	Kumano Basin, S of Japan	hydrate-bound gas	173	-40.5	Pape et al. (2014)
S of Pakistan							
12303	Nascent Ridge	Gravity Corer	Makran Accretionary Prism	hydrate-bound gas	6,463	-66.4	Fischer et al. (2013); Römer et al. (2012)
12316-3	Flare 2	Gravity Corer	Makran Accretionary Prism	hydrate-bound gas	3,632	-67.5	Fischer et al. (2013); Römer et al. (2012)
12316-4	Flare 2	Gravity Corer	Makran Accretionary Prism	hydrate-bound gas	6,173	-68.8	Fischer et al. (2013); Römer et al. (2012)
North Atlantic							
16807-2	Area 1	Gas Bubble Sampler	W of Spitsbergen	vent gas	6,363	-43.6	Sahling et al. (2014)
16823-2	Area 2	Gas Bubble Sampler	W of Spitsbergen	vent gas	7,497	-55.8	Sahling et al. (2014)
16823-5	Area 2	Gas Bubble Sampler	W of Spitsbergen	vent gas	7,418	-55.9	Sahling et al. (2014)
16833-2	Area 3	Gas Bubble Sampler	W of Spitsbergen	vent gas	7,748	-53.8	Sahling et al. (2014)
16833-3	Area 3	Gas Bubble Sampler	W of Spitsbergen	vent gas	8,385	-57.4	Sahling et al. (2014)
16848-2	Area 4	Gas Bubble Sampler	W of Spitsbergen	vent gas	9,028	-56.1	Sahling et al. (2014)
	Haakon Mosby MV	Gravity Corer	SW Barents Sea	hydrate-bound gas	1,215	n.det.	Pape et al. (2011)
	Haakon Mosby MV	Gravity Corer	SW Barents Sea	hydrate-bound gas	1,637	-63	Pape et al. (2011)

Table B-2: Methane hydrate data. $\delta^{13}\text{C}$ is in reference to PDB, δD is in reference to SMOW. $\delta^{13}\text{C}$ (B) indicates that this was measured in Bremen.

Sample/Site Description	$\delta^{13}\text{C}$	s.d.	δD	s.d.	$\Delta^{13}\text{CH}_3\text{D}$	s.d.	T(°C)	C1/C2	$\delta^{13}\text{C}$ (B)	Depth (cmbsf)
Baboon Hole	-73.73	0.001	-203.20	0.001	5.05	0.20	53	1,638	-69.3	-
Batumi Seep	-53.77	0.001	-232.50	0.001	4.97	0.34	57	4,178	-53.8	-
Batumi Seep	-54.30	0.001	-235.25	0.001	4.86	0.31	61	5,383	-53.2	-
Batumi Seep	-53.93	0.001	-234.58	0.001	4.93	0.16	58	4,631	-51.5	-
Bullseye Vent	-68.43	0.001	-163.99	0.001	7.21	0.45	-20	-	-	-
Bush Hill, Gulf of Mexico	-45.38	0.001	-174.90	0.001	3.67	0.42	106	-	-	-
Cascadia Margin, Oregon	-67.85	0.001	-170.86	0.001	5.57	0.44	25	-	-	-
Colkheti Seep	-50.03	0.001	-218.26	0.001	4.64	0.25	70	32	-49.5	-
Colkheti Seep	-49.92	0.001	-217.74	0.001	4.37	0.62	82	190	-48.1	-
Flare 2	-72.68	0.001	-216.00	0.001	4.62	0.36	71	3,632	-67.5	-
Flare 2	-72.90	0.001	-212.03	0.001	5.11	0.41	51	6,173	-68.8	-
Hegoland MV	-64.51	0.001	-240.30	0.001	3.27	0.28	143	3,054	-63.6	-
Hegoland MV	-63.62	0.001	-238.75	0.001	3.09	0.31	156	2,257	-62.8	-
Hydrate Hole	-74.03	0.001	-199.33	0.001	5.30	0.40	44	1,988	-70.6	-
Iberia Mound	-50.08	0.001	-239.92	0.001	4.50	0.41	76	2,090	-53.1	-
Kerch Flare	-72.44	0.001	-281.62	0.001	4.69	0.12	68	2,498	-66.9	-
Nascent Ridge	-72.14	0.001	-208.72	0.001	4.61	0.39	71	6,463	-66.4	-
Ordu Patch 2	-73.88	0.001	-247.91	0.001	5.48	0.26	38	3,131	-70.8	-
Ordu Patch 3	-74.05	0.001	-244.27	0.001	4.91	0.20	59	2,816	-70.4	-
Ordu Patch 5	-73.19	0.001	-240.81	0.001	5.29	0.19	45	2,335	-72.6	-
Ordu Patch 7	-73.29	0.001	-247.96	0.001	5.20	0.24	48	3,258	-71.3	-
Pechori Mound 1 cc	-49.79	0.001	-237.27	0.001	4.76	0.42	65	5	-48.8	-1
Pechori Mound 1/23 cm	-49.27	0.001	-233.92	0.001	5.06	0.97	53	not det	not det	0
Pechori Mound 5cc	-49.74	0.001	-239.15	0.001	4.83	0.35	62	not det	not det	-5
Pechori Mound 7cc	-52.60	0.001	-238.49	0.001	3.52	0.34	127	not det	not det	-7
Pechori Mound 9cc	-50.28	0.001	-237.15	0.001	4.21	0.33	90	not det	not det	-9
Pechori Mound, Black Sea	-48.61	0.001	-192.13	0.001	5.13	0.23	40	not det	not det	-
Pock A, S of Nigeria	-50.24	0.001	-161.32	0.001	4.24	0.39	77	8,443	-51	-
Pock C, S of Nigeria	-52.62	0.001	-160.57	0.001	4.71	0.19	56	6,467	-51.3	-
Pockmark A	-53.37	0.001	-193.58	0.001	4.90	0.18	59	8,443	-51	-
Pockmark C	-54.87	0.001	-193.66	0.001	5.37	0.51	42	6,467	-51.3	-
Poti Seep	-55.90	0.001	-235.06	0.001	4.83	0.38	62	4,153	-50.3	-
Vodyanitskii MV	-63.09	0.001	-235.06	0.001	2.74	0.12	184	2,018	-59.2	-

APPENDIX C

THE EFFECT OF THE ANAEROBIC OXIDATION OF METHANE ON CLUMPED METHANE ISOTOPOLOGUE ($^{13}\text{CH}_3\text{D}$) SYSTEMATICS

Motivation and Outlook

Clumped isotopologues (*e.g.*, $^{13}\text{CH}_3\text{D}$) of microbial methane from pure cultures of methanogens and quickly-generated microbial methane in environments such as swamps correspond to reportedly high methane equilibrium temperatures largely driven by kinetic processes (Douglas et al., 2017; Gruen et al., 2018, 2014; Stolper et al., 2013, 2014, 2015; Wang et al., 2015; Young et al., 2017). However, this is inconsistent with measurements of putatively microbial methane from environments such as methane hydrates (Wang et al., 2015, Chp. 2). It is hypothesized that this may be due to isotope exchange reactions, extremely slow microbial methanogenesis, or oxidation processes happening in the environment. However, isotope exchange with water is likely sluggish in these environments (Reeves et al., 2012). Methanogenesis occurring on geological timescales is difficult to measure. Moreover, the aerobic oxidation of methane by microbes actually pushes the clumped isotopologue signal even lower, farther from stochastic distribution in what is referred to as "anti-clumped" (Wang et al., 2016). So what is producing the near-equilibrium signals observed in oceanic sediments? One hypothesis is that other oxidation processes occurring here, such as the anaerobic oxidation of methane (AOM) may be acting on this methane reservoir. It is reported that most of the putatively microbial methane from Hydrate Ridge, Cascadia Margin, is actually removed by AOM (Boetius and Suess, 2004).

The goal of this work was to carry out culture experiments using incubations of sediments with active AOM metabolism (Holler et al., 2009; Yoshinaga et al., 2014; Wegener et al., 2016) to test the effect of AOM on the clumped isotope systematics of reservoirs of methane (Table C-1). The data in this appendix represent data that was sampled and measured but not incorporated into the main thesis chapters. It was found that in these AOM cultures, $\delta^{13}\text{C}$ and δD increased over the course of the methane consumption experiment, demonstrating kinetic isotope fractionation (Figures C-1 and C-2). The value of $\Delta^{13}\text{CH}_3\text{D}$ also increased, but in the direction of equilibrium values for the growth temperature (Figure C-3). Plotting the associated $^{13}\text{CH}_3\text{D}$ -based temperature illustrates this point, as over time apparent methane temperatures approach reasonable culture temperature (Figure C-4). In summary, AOM represents one explanation for the discrepancy between clumped isotopologue values observed in pure methanogen cultures compared with environmental samples. It is possible that AOM acts in the opposite direction from the aerobic oxidation of methane. This will be an exciting target of future work.

Acknowledgements

My sincerest thanks to Gunter Wegener, Martin Koenneke, Tobias Goldhammer, Heidi Taubner, Kai Hinrichs, and Shuhei Ono for supporting this work. DG, GW, MK, conceived

of this study. GW performed the culturing experiments. TG measured water isotopes. HT measured dissolved inorganic carbon concentrations and carbon isotopes. MK assisted with sample collection, analyses, and logistics in Bremen. DG measured all methane samples for clumped methane and conventional carbon and hydrogen isotopes.

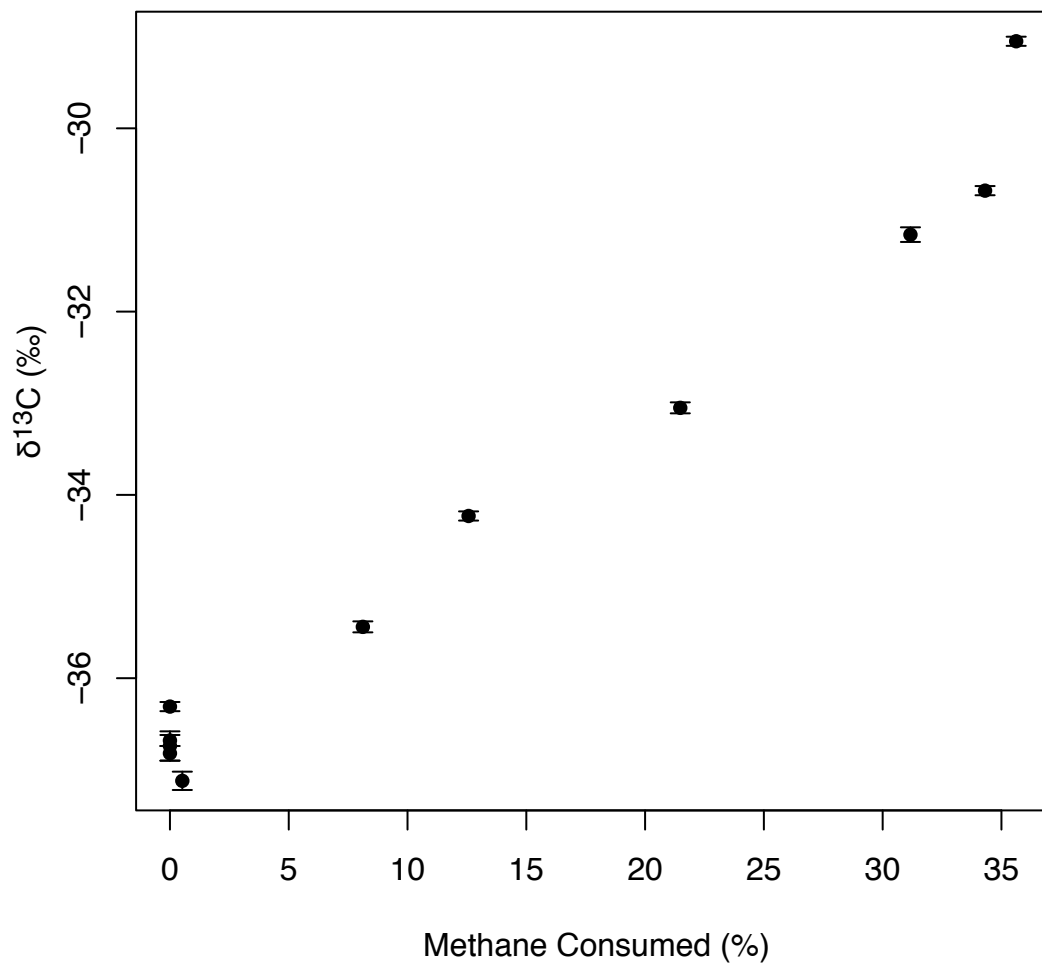


Figure C-1: $\delta^{13}\text{C}$ as a function of methane consumed via AOM.

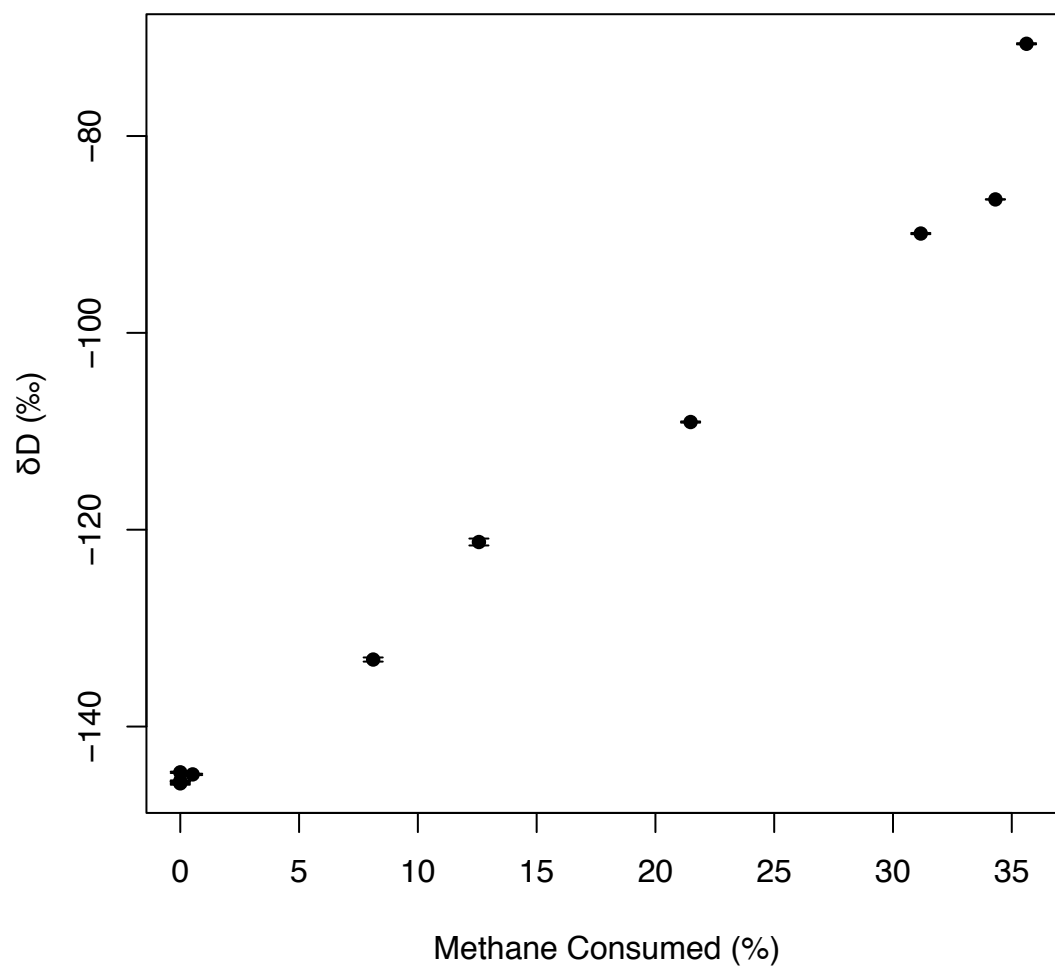


Figure C-2: δD as a function of methane consumed via AOM.

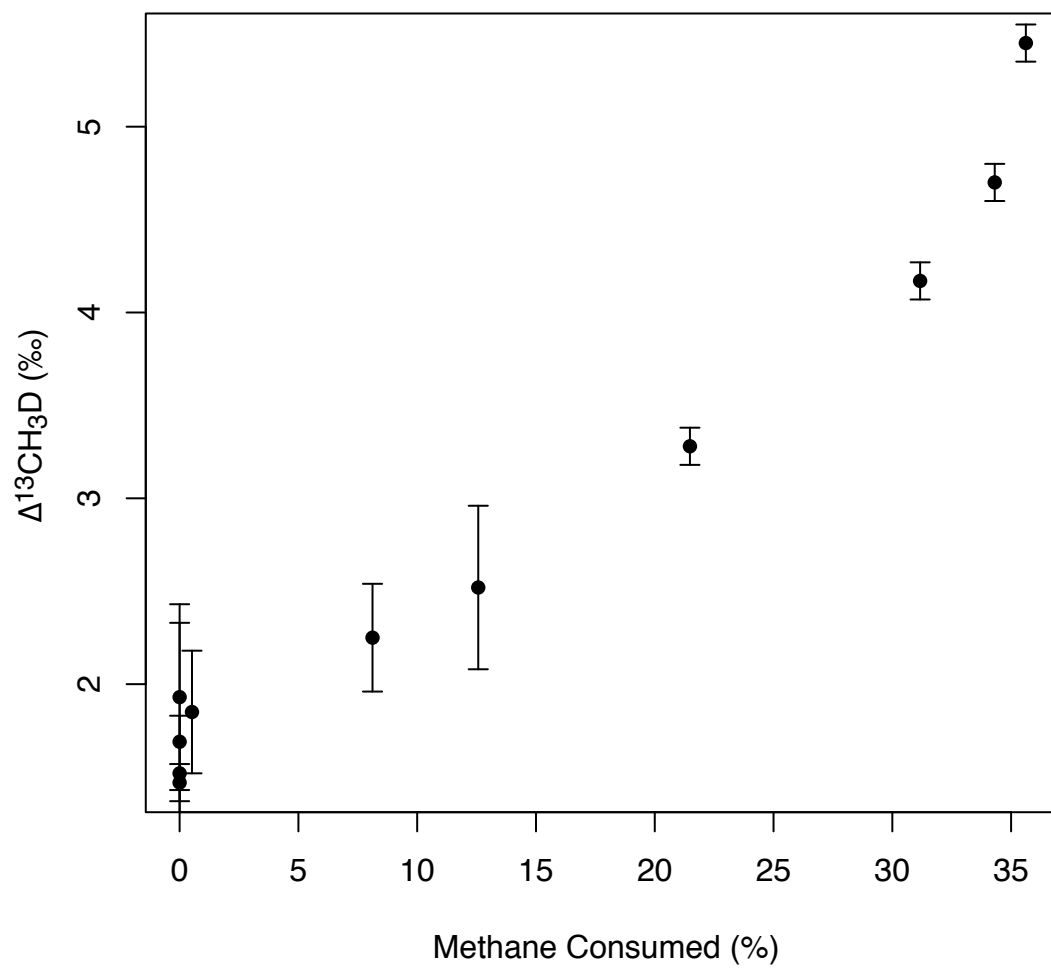


Figure C-3: $\Delta^{13}\text{CH}_3\text{D}$ as a function of methane consumed via AOM.

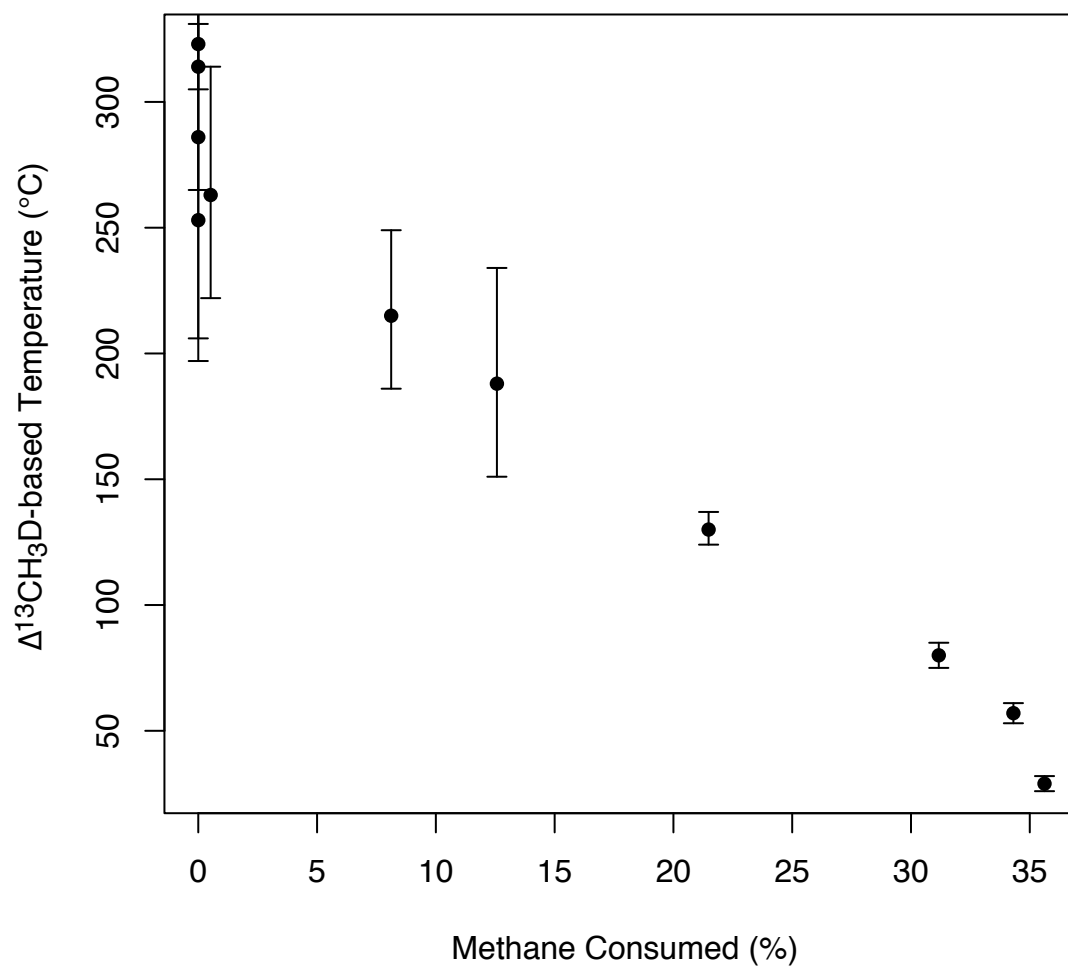


Figure C-4: $\Delta^{13}\text{CH}_3\text{D}$ -based temperature as a function of methane consumed via AOM.

Table C-1: Methane isotope composition over the course of the methane consumption experiment.

CH ₄ consumed (%)	$\delta^{13}\text{C}$	δD	$\Delta^{13}\text{CH}_3\text{D}$	T(°C)
0.00	-36.74 ± 0.16 ‰	-145.77 ± 0.14 ‰	1.69 ± 0.64 ‰	286 +132/-80 °C
0.52	-37.12 ± 0.10 ‰	-144.86 ± 0.07 ‰	1.85 ± 0.33 ‰	263 +51/-41 °C
0.00	-36.82 ± 0.08 ‰	-145.70 ± 0.09 ‰	1.93 ± 0.50 ‰	253 +78/-56 °C
8.12	-35.44 ± 0.06 ‰	-133.19 ± 0.21 ‰	2.25 ± 0.29 ‰	215 +34/-29 °C
12.57	-34.23 ± 0.05 ‰	-121.25 ± 0.35 ‰	2.52 ± 0.44 ‰	188 +46/-37 °C
21.48	-33.05 ± 0.06 ‰	-109.06 ± 0.06 ‰	3.28 ± 0.10 ‰	130 +7/-6 °C
0.00	-36.68 ± 0.06 ‰	-145.52 ± 0.05 ‰	1.47 ± 0.10 ‰	323 +19/-18 °C
31.17	-31.16 ± 0.08 ‰	-89.91 ± 0.06 ‰	4.17 ± 0.10 ‰	80 +5/-5 °C
34.31	-30.68 ± 0.05 ‰	-86.44 ± 0.01 ‰	4.70 ± 0.10 ‰	57 +4/-4 °C
35.62	-29.05 ± 0.05 ‰	-70.63 ± 0.06 ‰	5.45 ± 0.10 ‰	29 +3/-3 °C
0.00	-36.31 ± 0.05 ‰	-144.65 ± 0.07 ‰	1.52 ± 0.31 ‰	314 +63/-49 °C

APPENDIX D

EVOLUTION OF THE METHYL-CORRINOID PATHWAY OF METHANOGENESIS AND POSSIBLE IMPLICATIONS FOR BIOGEOCHEMICAL CYCLING IN PROTEROZOIC AND PHANEROZOIC OCEANS

Supplemental Figures

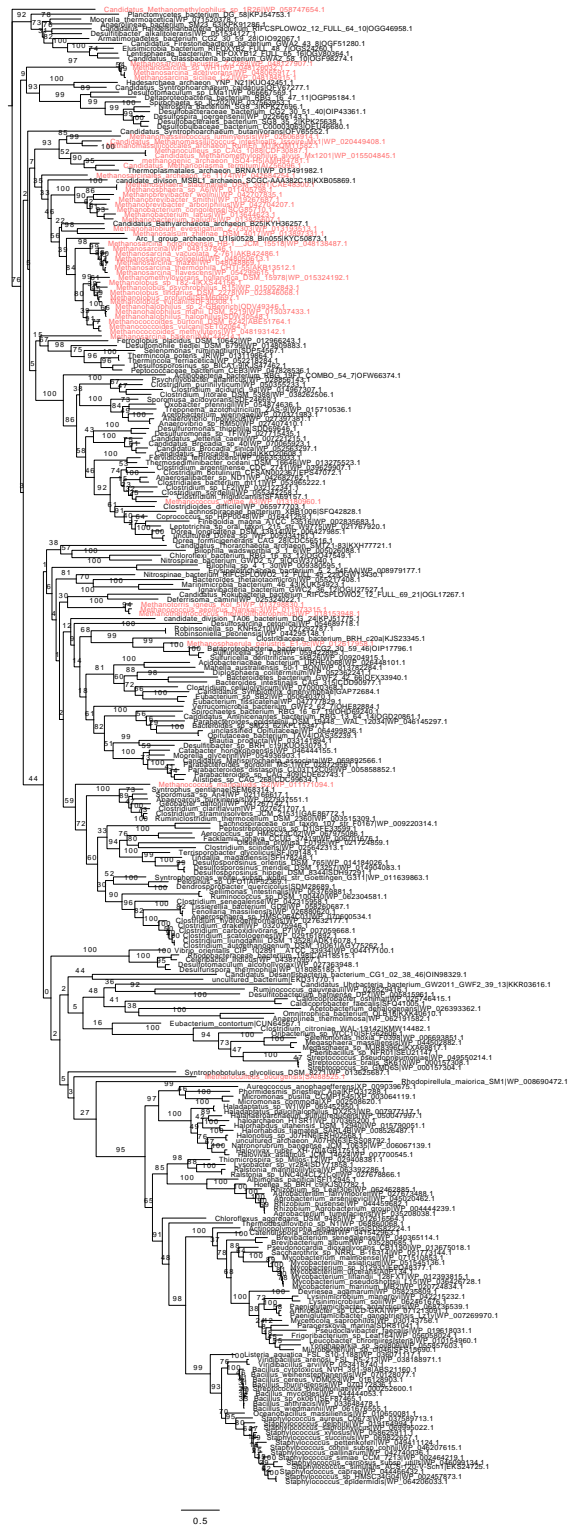


Figure D-1: Gene tree of *cmtA*. Red indicates putatively methanogenic taxa.



182



Figure D-3: Gene tree of *mtaB*. Red indicates putatively methanogenic taxa.



Figure D-4: Gene tree of mtbA. Red indicates putatively methanogenic taxa.

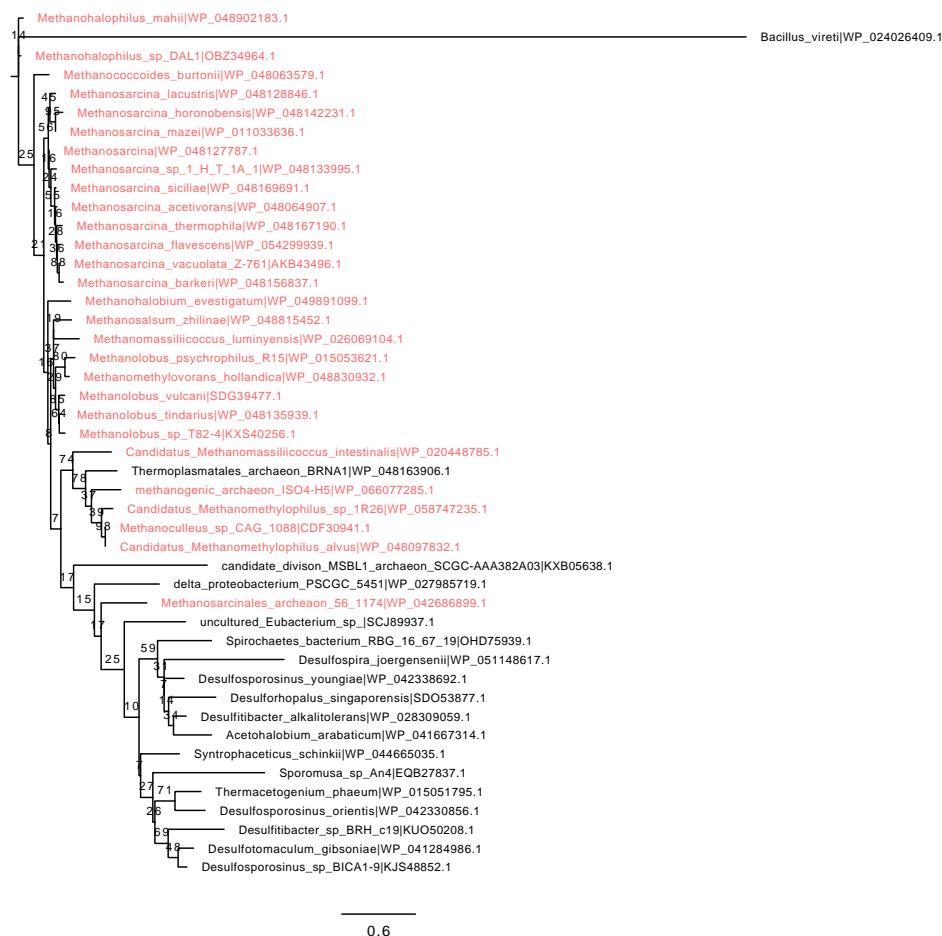


Figure D-5: Gene tree of *mtbB*. Red indicates putatively methanogenic taxa.

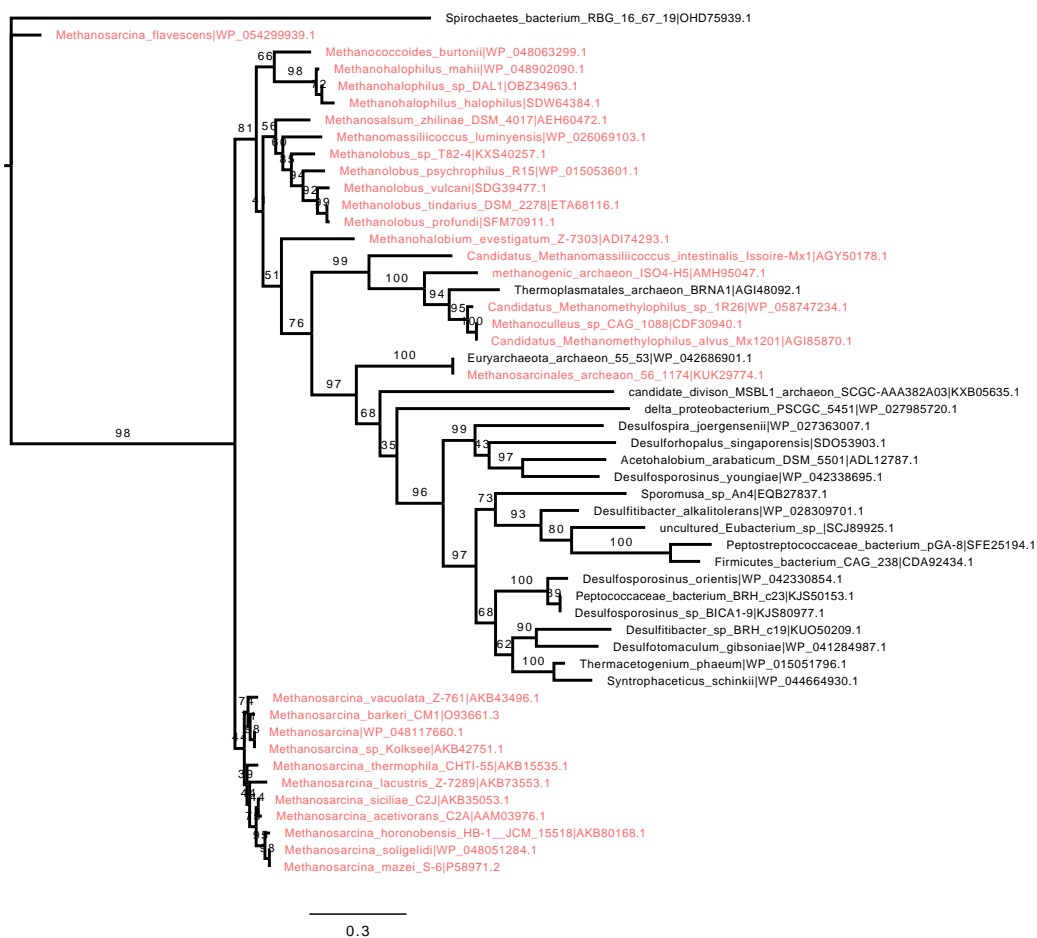


Figure D-6: Gene tree of mtbB1. Red indicates putatively methanogenic taxa.



Figure D-7: Gene tree of *mtbC*. Red indicates putatively methanogenic taxa.



Figure D-8: Gene tree of mtmB1. Red indicates putatively methanogenic taxa.



189



Figure D-10: Gene tree of *mtsB*. Red indicates putatively methanogenic taxa.



191

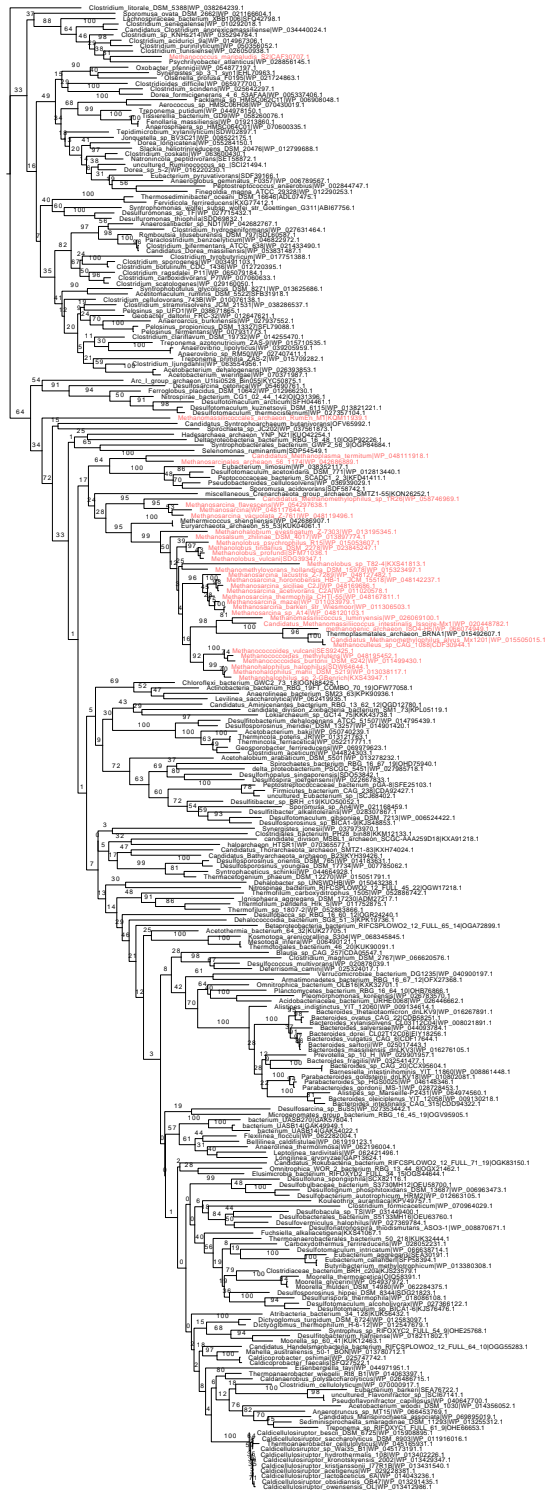


Figure D-12: Gene tree of *mttC*. Red indicates putatively methanogenic taxa.

APPENDIX E

SUPPLEMENTARY MATERIAL FOR CHITINASES: A STANDARD CANDLE FOR DATING MICROBIAL LINEAGES (CHAPTER 4)



Figure E-1: Expanded RAxML gene tree with tip labels and bootstrap support values.

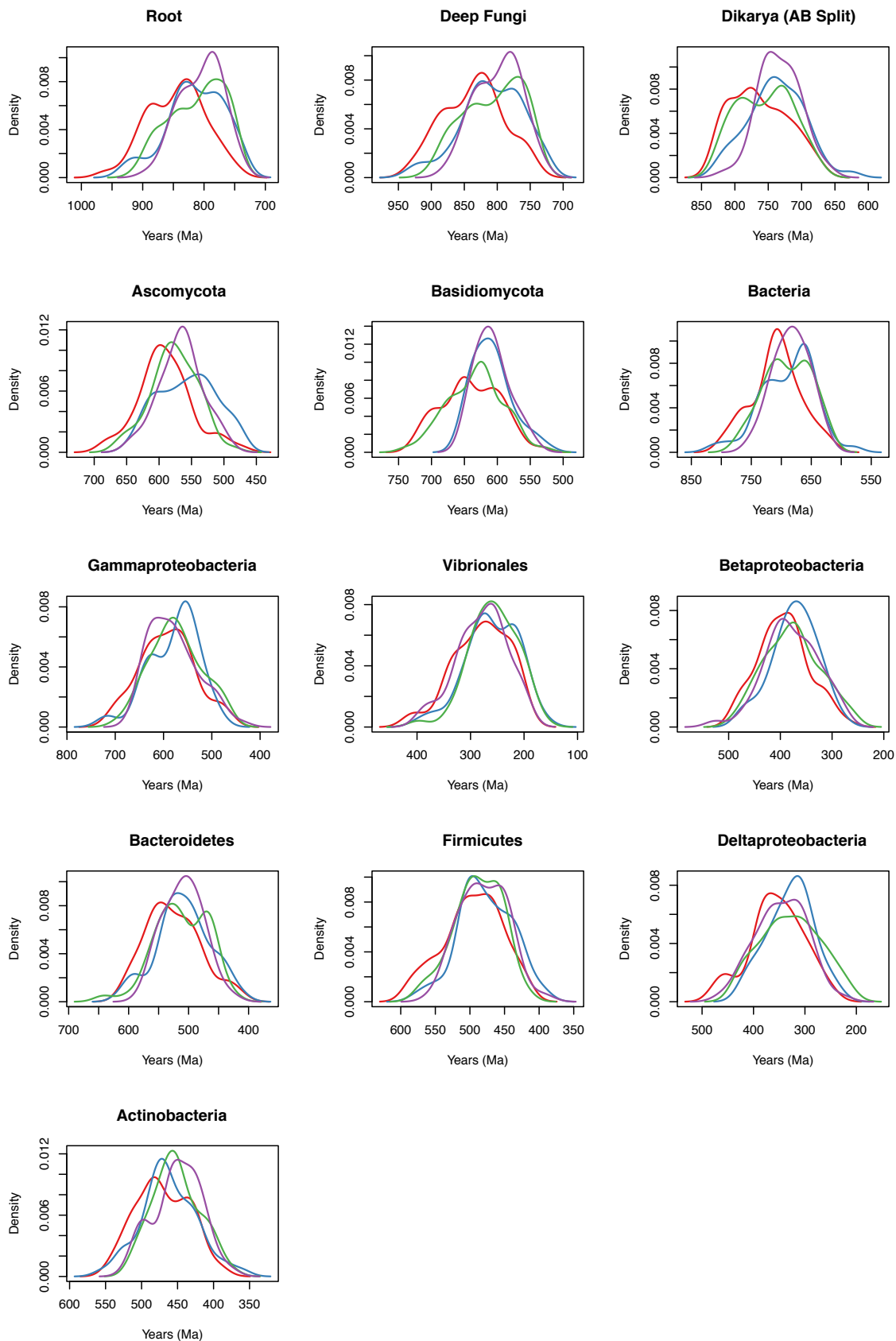


Figure E-2: Prior date distributions across 11 nodes under the four calibration setups.

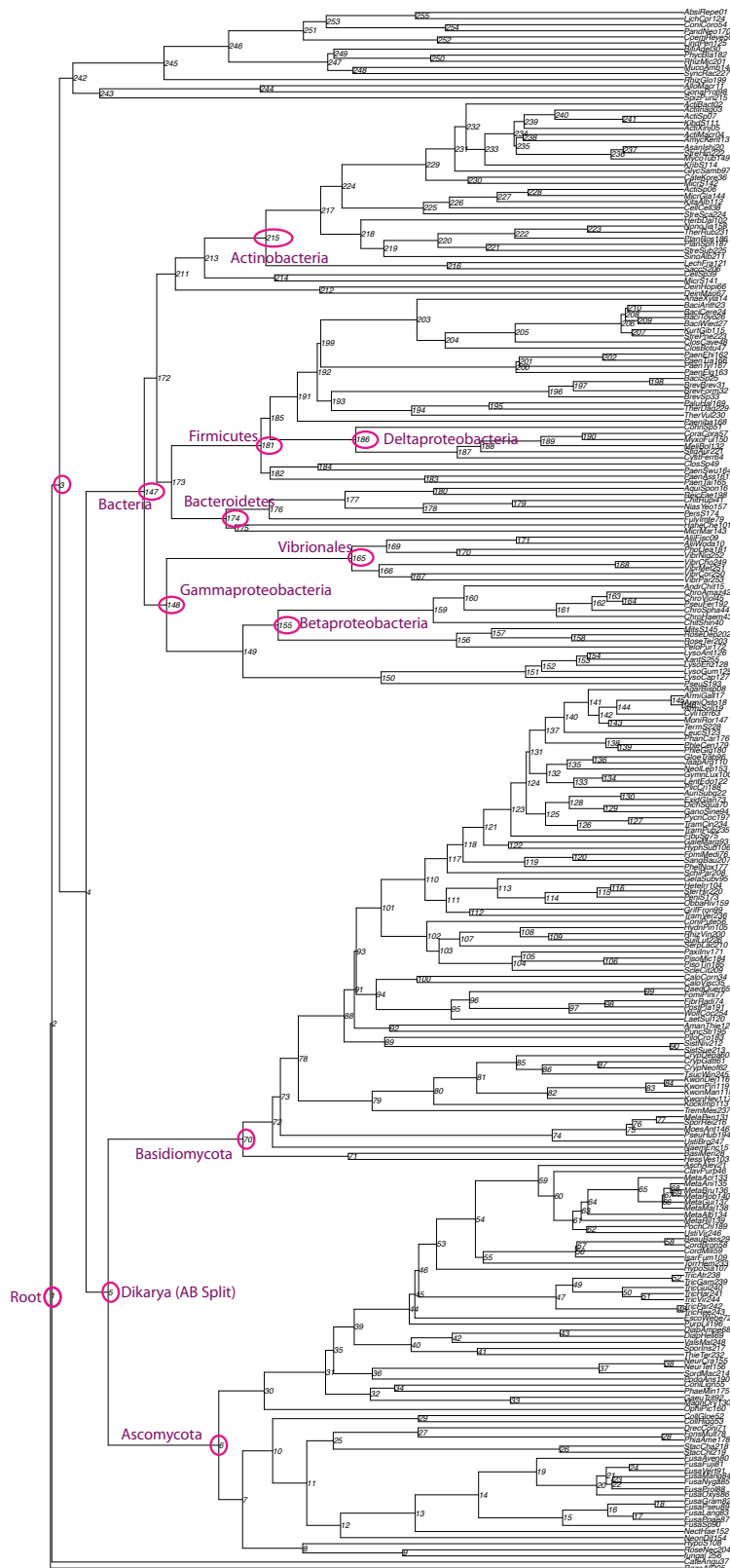


Figure E-3: Chronogram with corresponding node numbers used in analysis. Clades are annotated on corresponding nodes.

Table E-1: Taxa in this study. Environment refers to the environment that these substrates operate in or where the organism was sampled from according to NCBI. Aquatic represents marine environments. Terrestrial refers to land which includes shallow freshwater ponds/swamp type sample locations.

Taxon Name Protein ID	Clade	Taxonomy	Environment
Actinobacteria_bacterium_13_2_20CM_2_71_6 OLB77094.1	bacteria	Actinobacteria	T
Actinokineospora_inagensis WP_026423506.1	bacteria	Actinobacteria	T
Actinomadura_macra WP_067464737.1	bacteria	Actinobacteria	T
Actinophytocola_xinjiangensis WP_075132815.1	bacteria	Actinobacteria	T
Actinosporangium_sp_NRRL_B_3428 WP_052581227.1	bacteria	Actinobacteria	T
Actinosynnema_sp_ALI_1_44 WP_076987693.1	bacteria	Actinobacteria	T
Amycolatopsis_kentuckyensis WP_086844859.1	bacteria	Actinobacteria	T
Asanoa_ishikariensis WP_090795177.1	bacteria	Actinobacteria	T
Catelliglobospora_koreensis WP_026207975.1	bacteria	Actinobacteria	T
Cellulosimicrobium_cellulans WP_087471431.1	bacteria	Actinobacteria	T
Glycomyces_sambucus WP_091045549.1	bacteria	Actinobacteria	T
Herbidospora_daliensis WP_062439410.1	bacteria	Actinobacteria	T
Kibdelosporangium_sp_MJ126_NF4 WP_042195239.1	bacteria	Actinobacteria	T
Kitasatospora_albolonga WP_084750020.1	bacteria	Actinobacteria	T
Kribbella_sp_ALI_6_A WP_077015618.1	bacteria	Actinobacteria	T
Lechevalieria_fradiae WP_090044473.1	bacteria	Actinobacteria	T
Micromonospora_sp_CB01531 WP_073839717.1	bacteria	Actinobacteria	T
Microtetraspora_glauca WP_030497465.1	bacteria	Actinobacteria	T
Mycobacterium_tuberculosis CNE32205.1	bacteria	Actinobacteria	T
Nonomuraea_jiangxiensis WP_090946548.1	bacteria	Actinobacteria	T
Planobispora_rosea WP_068921834.1	bacteria	Actinobacteria	T
Planomonospora_sphaerica WP_068895167.1	bacteria	Actinobacteria	T
Saccharothrix_sp_NRRL_B_16348 WP_053716888.1	bacteria	Actinobacteria	T
Sinosporangium_album WP_093169663.1	bacteria	Actinobacteria	T
Streptoalloteichus_hindustanus WP_073483646.1	bacteria	Actinobacteria	T
Streptomyces_scabrisporus WP_020551069.1	bacteria	Actinobacteria	T
Streptosporangium_subroseum WP_089208112.1	bacteria	Actinobacteria	T
Thermoactinospira_rubra WP_084964699.1	bacteria	Actinobacteria	T
Aquimarina_spongiae SHI61232.1	bacteria	Bacteroidetes	A
Chitinophaga_rupis WP_089906523.1	bacteria	Bacteroidetes	T
Fulvivirga_imtechensis WP_083867410.1	bacteria	Bacteroidetes	A
Microscilla_marina_ATCC_23134 EAY30869.1	bacteria	Bacteroidetes	A
Niastella_yeongjuensis WP_081197334.1	bacteria	Bacteroidetes	T
Persicobacter_sp_JZB09 WP_060687981.1	bacteria	Bacteroidetes	A
Reichenbachiella_faecimaris WP_084370698.1	bacteria	Bacteroidetes	A
Andrieprevotia_chitinilytica WP_084187288.1	bacteria	Betaproteobacteria	T
Chitiniphilus_shinanonensis WP_018748575.1	bacteria	Betaproteobacteria	A
Chromobacterium_amazonense WP_071108352.1	bacteria	Betaproteobacteria	A
Chromobacterium_haemolyticum WP_081574973.1	bacteria	Betaproteobacteria	T
Chromobacterium_sphagni WP_071113284.1	bacteria	Betaproteobacteria	T

Continued on next page

Table E-1 – continued from previous page

Taxon Name Protein ID	Clade	Taxonomy	Environment
Chromobacterium_violaceum WP_043615248.1	bacteria	Betaproteobacteria	T
Mitsuaria_sp_7 WP_082938754.1	bacteria	Betaproteobacteria	T
Pelomonas_puraquae WP_088484787.1	bacteria	Betaproteobacteria	T
Pseudogulbenkiania_ferroxidans WP_031296507.1	bacteria	Betaproteobacteria	T
Roseateles_depolymerans WP_083526021.1	bacteria	Betaproteobacteria	T
Roseateles_terrae WP_088453654.1	bacteria	Betaproteobacteria	T
Deinococcus_hopiensis_KR_140 WP_084045460.1	bacteria	Deinococcus	T
Deinococcus_maricopenensis_DSM_21211 WP_013558263.1	bacteria	Deinococcus	T
Cohnella_sp_CIP_111063 WP_094044771.1	bacteria	Deltaproteobacteria	T
Corallococcus_coralloides_DSM_2259 WP_014399530.1	bacteria	Deltaproteobacteria	T
Cystobacter_ferrugineus WP_084736787.1	bacteria	Deltaproteobacteria	T
Melittangium_boletus_DSM_14713 WP_095979189.1	bacteria	Deltaproteobacteria	T
Myxococcus_fulvus WP_046715376.1	bacteria	Deltaproteobacteria	T
Stigmatella_aurantiaca_DW4_3_1 WP_013376730.1	bacteria	Deltaproteobacteria	T
Anaerocolumna_xylanovorans_DSM_12503 WP_073590321.1	bacteria	Firmicutes	T
Bacillus_anthraxis_str_H9401 WP_014654726.1	bacteria	Firmicutes	T
Bacillus_cereus_03BB108 WP_001994684.1	bacteria	Firmicutes	T
Bacillus_sp_FJAT_27238 WP_016742699.1	bacteria	Firmicutes	NR
Bacillus_toyonensis PAW47140.1	bacteria	Firmicutes	T
Bacillus_wiedmannii WP_098079181.1	bacteria	Firmicutes	T
Brevibacillus_brevis WP_087349255.1	bacteria	Firmicutes	T
Brevibacillus_formosus WP_047074631.1	bacteria	Firmicutes	T
Brevibacillus_sp_Leaf182 WP_056491684.1	bacteria	Firmicutes	T
Clostridium_botulinum_B_str_Osaka05 WP_073860743.1	bacteria	Firmicutes	T
Clostridium_cavendishii_DSM_21758 SHI69443.1	bacteria	Firmicutes	T
Clostridium_sp_ND2 WP_084764479.1	bacteria	Firmicutes	NR
Kurthia_gibsonii AFI72779.1	bacteria	Firmicutes	T
Paenibacillus_assamensis WP_051217339.1	bacteria	Firmicutes	T
Paenibacillus_ehimensis WP_025852116.1	bacteria	Firmicutes	T
Paenibacillus_elgii WP_063184838.1	bacteria	Firmicutes	T
Paenibacillus_swuensis WP_068603214.1	bacteria	Firmicutes	T
Paenibacillus_taiwanensis WP_051287538.1	bacteria	Firmicutes	T
Paenibacillus_tianmuensis WP_090673949.1	bacteria	Firmicutes	T
Paenibacillus_tyrfis WP_036687954.1	bacteria	Firmicutes	T
Paenibacillus WP_081717951.1	bacteria	Firmicutes	T
Paludifilum_halophilum WP_094265818.1	bacteria	Firmicutes	T
Streptococcus_pneumoniae CKG26627.1	bacteria	Firmicutes	T
Thermoactinomyces_daqus WP_081944003.1	bacteria	Firmicutes	T
Thermoactinomyces_vulgaris WP_022737757.1	bacteria	Firmicutes	T
Hahella_chejuensis_KCTC_2396 WP_011394837.1	bacteria	Gammaproteobacteria	A
Aliivibrio_fischeri WP_065597214.1	bacteria	Gammaproteobacteria	A
Aliivibrio_wodanis WP_061013584.1	bacteria	Gammaproteobacteria	A
Cellvibrio_sp_pealriver WP_049631752.1	bacteria	Gammaproteobacteria	A
Lysobacter_antibioticus WP_079248132.1	bacteria	Gammaproteobacteria	T

Continued on next page

Table E-1 – continued from previous page

Taxon Name Protein ID	Clade	Taxonomy	Environment
Lysobacter_capsici ALN87908.1	bacteria	Gammaproteobacteria	T
Lysobacter_enzymogenes WP_082644261.1	bacteria	Gammaproteobacteria	T
Lysobacter_gummosus ALN93704.1	bacteria	Gammaproteobacteria	T
Microbulbifer_sp_HZ11 WP_081847999.1	bacteria	Gammaproteobacteria	A
Photobacterium_jeanii WP_068327670.1	bacteria	Gammaproteobacteria	A
Pseudoxanthomonas_sp_CF125 SDQ75949.1	bacteria	Gammaproteobacteria	T
Vibrio_cholerae WP_095481188.1	bacteria	Gammaproteobacteria	A
Vibrio_coralliilyticus WP_038509898.1	bacteria	Gammaproteobacteria	A
Vibrio_metoecus KQA22300.1	bacteria	Gammaproteobacteria	T
Vibrio_nigripulchritudo WP_022594672.1	bacteria	Gammaproteobacteria	A
Vibrio_parahaemolyticus WP_031855765.1	bacteria	Gammaproteobacteria	A
Xanthomonas_sp_AK BAA36460.1	bacteria	Gammaproteobacteria	T
Aschersonia_aleyrodinis_RCEF_2490 KZZ88173.1	fungi	Ascomycota	
Beauveria_bassiana_ARSEF_2860 XP_008603412.1	fungi	Ascomycota	
Claviceps_purpurea_20_1 CCE29524.1	fungi	Ascomycota	
Colletotrichum_gloeosporioides_Cg_14 EQB55915.1	fungi	Ascomycota	
Colletotrichum_higginsianum_IMI_349063 XP_018160050.1	fungi	Ascomycota	
Coniochaeta_ligniaria_NRRL_30616 OIW32936.1	fungi	Ascomycota	
Cordyceps_brongniartii_RCEF_3172 OAA33915.1	fungi	Ascomycota	
Cordyceps_militaris ATY67196.1	fungi	Ascomycota	
Diaporthe_ampelina KKY33732.1	fungi	Ascomycota	
Diaporthe_helianthi POS73161.1	fungi	Ascomycota	
Drechmeria_coniospora KYK61674.1	fungi	Ascomycota	
Escovopsis_weberi KOS21945.1	fungi	Ascomycota	
Fonsecaea_multimorphosa_CBS_102226 XP_016628205.1	fungi	Ascomycota	
fungus_sp_No_14919 GAW12378.1	fungi	Ascomycota	
Fusarium_avenaceum KIL91362.1	fungi	Ascomycota	
Fusarium_fujikuroi_IMI_58289 XP_023429066.1	fungi	Ascomycota	
Fusarium_graminearum_PH_1 XP_011324582.1	fungi	Ascomycota	
Fusarium_langsethiae KPA45507.1	fungi	Ascomycota	
Fusarium_mangiferae CVK87365.1	fungi	Ascomycota	
Fusarium_nygamai PNP84899.1	fungi	Ascomycota	
Fusarium_oxysporum_FOSC_3_a EWZ02462.1	fungi	Ascomycota	
Fusarium_poae OBS21759.1	fungi	Ascomycota	
Fusarium_proliferatum CVK87747.1	fungi	Ascomycota	
Fusarium_pseudograminearum_CS3096 XP_009254403.1	fungi	Ascomycota	
Fusarium_sp_FIESC_5_CS3069 CEG04533.1	fungi	Ascomycota	
Fusarium_verticillioides_7600 XP_018749091.1	fungi	Ascomycota	
Gaeumannomyces_tritici_R3_111a_1 XP_009219406.1	fungi	Ascomycota	
Hypocrella_siamensis ALI93553.1	fungi	Ascomycota	
Hypoxylon_sp_CI_4A OTB05187.1	fungi	Ascomycota	
Isaria_fumosorosea_ARSEF_2679 XP_018702722.1	fungi	Ascomycota	
Magnaporthe_oryzae_70_15 XP_003714897.1	fungi	Ascomycota	
Metarhizium_acridum_CQMa_102 XP_007815036.1	fungi	Ascomycota	

Continued on next page

Table E-1 – continued from previous page

Taxon Name Protein ID	Clade	Taxonomy	Environment
Metarhizium_album_ARSEF_1941 KHN93916.1	fungi	Ascomycota	
Metarhizium_anisopliae_BRIP_53293 KJK76262.1	fungi	Ascomycota	
Metarhizium_brunneum_ARSEF_3297 XP_014539677.1	fungi	Ascomycota	
Metarhizium_guizhouense_ARSEF_977 KID83374.1	fungi	Ascomycota	
Metarhizium_majus_ARSEF_297 XP_014573577.1	fungi	Ascomycota	
Metarhizium_rileyi_RCEF_4871 OAA44119.1	fungi	Ascomycota	
Metarhizium_robertsii_ARSEF_23 XP_007823947.1	fungi	Ascomycota	
Nectria_haematococca_mpVI_77_13_4 XP_003050159.1	fungi	Ascomycota	
Neonectria_ditissima KPM38261.1	fungi	Ascomycota	
Neurospora_crassa_OR74A XP_011395327.1	fungi	Ascomycota	
Neurospora_tetrasperma_FGSC_2508 XP_009853080.1	fungi	Ascomycota	
Ophiostoma_piceae_UAMH_11346 EPE05896.1	fungi	Ascomycota	
Phaeoacremonium_minimum_UCRPA7 XP_007917527.1	fungi	Ascomycota	
Phialophora_americana KIW68338.1	fungi	Ascomycota	
Pochonia_chlamydosporia_170 XP_018136544.1	fungi	Ascomycota	
Podospora_anserina_S_mat_ XP_001904165.1	fungi	Ascomycota	
Purpureocillium_lilacinum XP_018179211.1	fungi	Ascomycota	
Rosellinia_necatrix GAP83950.1	fungi	Ascomycota	
Sordaria_macrospora_k_hell XP_003348594.1	fungi	Ascomycota	
Sporothrix_insectorum_RCEF_264 OAA58430.1	fungi	Ascomycota	
Stachybotrys_chartarum_IBT_40293 KFA46477.1	fungi	Ascomycota	
Stachybotrys_chlorohalonata_IBT_40285 KFA60985.1	fungi	Ascomycota	
Thielavia_terrestris_NRR_L_8126 XP_003654697.1	fungi	Ascomycota	
Torrubiella_hemipterigena CEJ94273.1	fungi	Ascomycota	
Trichoderma_atroviride_IMI_206040 XP_013945238.1	fungi	Ascomycota	
Trichoderma_gamsii XP_018659094.1	fungi	Ascomycota	
Trichoderma_guizhouense OPB46374.1	fungi	Ascomycota	
Trichoderma_harzianum PNP57499.1	fungi	Ascomycota	
Trichoderma_parareesei OTA01922.1	fungi	Ascomycota	
Trichoderma_reesei CAZ16624.1	fungi	Ascomycota	
Trichoderma_virens_Gv29_8 XP_013952675.1	fungi	Ascomycota	
Ustilaginoidea_virens KDB16030.1	fungi	Ascomycota	
Valsa_mali KUI66287.1	fungi	Ascomycota	
Agaricus_bisporus_var_burnettii_JB137_S8 XP_007327593.1	fungi	Basidiomycota	
Amanita_thiersii_Skay4041 PFH50181.1	fungi	Basidiomycota	
Armillaria_gallica PBK99462.1	fungi	Basidiomycota	
Armillaria_ostoyae SJJL03806.1	fungi	Basidiomycota	
Armillaria_solidipes PBK72370.1	fungi	Basidiomycota	
Auricularia_subglabra_TFB_10046_SS5 XP_007352083.1	fungi	Basidiomycota	
Calocera_cornea_HHB12733 KZT52238.1	fungi	Basidiomycota	
Calocera_viscosa_TUFC12733 KZP01089.1	fungi	Basidiomycota	
Coniophora puteana_RWD_64_598_SS2 XP_007770490.1	fungi	Basidiomycota	
Cryptococcus_depauperatus_CBS_7841 ODN91132.1	fungi	Basidiomycota	
Cryptococcus_gattii_E566 KIY34695.1	fungi	Basidiomycota	

Continued on next page

Table E-1 – continued from previous page

Taxon Name Protein ID	Clade	Taxonomy	Environment
Cryptococcus_neoformans_var_grubii OWZ72202.1	fungi	Basidiomycota	
Cylindrobasidium_torrendii_FP15055_ss_10 KIY62608.1	fungi	Basidiomycota	
Daedalea_quercina_L_15889 KZT71956.1	fungi	Basidiomycota	
Dichomitus_squalens_LYAD_421_SS1 XP_007362933.1	fungi	Basidiomycota	
Exidia_glandulosa_HHB12029 KZV81950.1	fungi	Basidiomycota	
Fibroporia_radiculosa XP_012181105.1	fungi	Basidiomycota	
Fibularhizoctonia_sp_CBS_109695 KZP14692.1	fungi	Basidiomycota	
Fomitiporia_mediterranea_MF3_22 XP_007261858.1	fungi	Basidiomycota	
Fomitopsis_pinicola_FP_58527_SS1 EPT05438.1	fungi	Basidiomycota	
Galerina_marginata_CBS_339_88 KDR80726.1	fungi	Basidiomycota	
Ganoderma_sinense_ZZ0214_1 PIL36523.1	fungi	Basidiomycota	
Gelatoporia_subvermispora_B EMD33176.1	fungi	Basidiomycota	
Gloeophyllum_trabeum_ATCC_11539 XP_007865094.1	fungi	Basidiomycota	
Grifola_frondosa OBZ73901.1	fungi	Basidiomycota	
Gymnopus_luxurians_FD_317_M1 KIK68276.1	fungi	Basidiomycota	
Heterobasidion_irregulare_TC_32_1 XP_009549140.1	fungi	Basidiomycota	
Hydnomerulius_pinastri_MD_312 KIJ57853.1	fungi	Basidiomycota	
Hypholoma_sublateritium_FD_334_SS_4 KJA28076.1	fungi	Basidiomycota	
Jaapia_argillacea_MUCL_33604 KDQ61715.1	fungi	Basidiomycota	
Kockovaella_imperatae XP_021871935.1	fungi	Basidiomycota	
Kwoniella_dejecticola_CBS_10117 XP_018260439.1	fungi	Basidiomycota	
Kwoniella_heveanensis_CBS_569 OCF46041.1	fungi	Basidiomycota	
Kwoniella_mangroviensis_CBS_8886 XP_019003580.1	fungi	Basidiomycota	
Kwoniella_pini_CBS_10737 XP_019009183.1	fungi	Basidiomycota	
Laetiporus_sulphureus_93_53 KZT09935.1	fungi	Basidiomycota	
Lentinula_edodes GAW00447.1	fungi	Basidiomycota	
Leucoagaricus_sp_SymC_cos KXN83991.1	fungi	Basidiomycota	
Melanopsichium_pennsylvanicum_4 CDI53624.1	fungi	Basidiomycota	
Moesziomyces_antarcticus XP_014657252.1	fungi	Basidiomycota	
Moniliophthora_roreri_MCA_2997 XP_007844272.1	fungi	Basidiomycota	
Naematelia_encephala ORY26060.1	fungi	Basidiomycota	
Neolentinus_lepideus_HHB14362_ss_1 KZT18638.1	fungi	Basidiomycota	
Obba_rivulosa OCH85900.1	fungi	Basidiomycota	
Paxillus_involutus_ATCC_200175 KIJ20918.1	fungi	Basidiomycota	
Peniophora_sp_CONT KZV61339.1	fungi	Basidiomycota	
Phanerochaete_carnosa_HHB_10118_sp XP_007400730.1	fungi	Basidiomycota	
Phellinus_noxius PAV18116.1	fungi	Basidiomycota	
Phlebia_centrifuga OKY58257.1	fungi	Basidiomycota	
Phlebiopsis_gigantea_11061_1_CR5_6 KIP08831.1	fungi	Basidiomycota	
Piloderma_croceum_F_1598 KIM83828.1	fungi	Basidiomycota	
Pisolithus_microcarpus_441 KIK23049.1	fungi	Basidiomycota	
Pisolithus_tinctorius_Marx_270 KIO10841.1	fungi	Basidiomycota	
Plicaturopsis_crispa_FD_325_SS_3 KII93375.1	fungi	Basidiomycota	
Postia_placenta_Mad_698_R XP_002475036.1	fungi	Basidiomycota	

Continued on next page

Table E-1 – continued from previous page

Taxon Name Protein ID	Clade	Taxonomy	Environment
Pseudozyma_hubeiensis_SY62 XP_012187014.1	fungi	Basidiomycota	
Punctularia_strigosozonata_HHB_11173_SS5 XP_007386461.1	fungi	Basidiomycota	
Pycnoporus_coccineus_BRFM310 OSD02937.1	fungi	Basidiomycota	
Rhizopogon_vinicolor_AM_OR11_026 OAX37953.1	fungi	Basidiomycota	
Sanghuangporus_baumii OCB91529.1	fungi	Basidiomycota	
Schizopora_paradoxa KLO18244.1	fungi	Basidiomycota	
Scleroderma_citrinum_Foug_A KIM58986.1	fungi	Basidiomycota	
Serpula_lacrymans_var_lacrymans_S7_9 XP_007320926.1	fungi	Basidiomycota	
Sistotremastrum_niveocreum_HHB9708 KZS88861.1	fungi	Basidiomycota	
Sistotremastrum_suecicum_HHB10207_ss_3 KZT36940.1	fungi	Basidiomycota	
Sporisorium_reilianum_f_sp_reilianum SJX62990.1	fungi	Basidiomycota	
Stereum_hirsutum_FP_91666_SS1 XP_007304512.1	fungi	Basidiomycota	
Suillus_luteus_UH_Slu_Lm8_n1 KIK48287.1	fungi	Basidiomycota	
Termitomyces_sp_J132 KNZ74502.1	fungi	Basidiomycota	
Trametes_cinnabarina CDO78089.1	fungi	Basidiomycota	
Trametes_pubescens OJT02884.1	fungi	Basidiomycota	
Trametes_versicolor_FP_101664_SS1 XP_008038395.1	fungi	Basidiomycota	
Tremella_mesenterica_DSM_1558 XP_007003148.1	fungi	Basidiomycota	
Tsuchiyaea_wingfieldii_CBS_7118 XP_019032869.1	fungi	Basidiomycota	
Ustilago_bromivora SAM82108.1	fungi	Basidiomycota	
Wolfiporia_cocos_MD_104_SS10 PCH44331.1	fungi	Basidiomycota	
Basidiobolus_meristosporus_CBS_931_73 ORX91522.1	fungi	Basidiomycota	
Hesseltinella_vesiculosa ORX58941.1	fungi	Basidiomycota	
Allomyces_macrogyneus_ATCC_38327 KNE60875.1	fungi	Blastocladiomycota	
Catenaria_anguillulae_PL171 ORZ38929.1	fungi	Blastocladiomycota	
Gonapodya_prolifera_JEL478 KXS16685.1	fungi	Chytridiomycota	
Rhizoclostridium_globosum ORY41636.1	fungi	Chytridiomycota	
Spizellomyces_punctatus_DAOM_BR117 XP_016608094.1	fungi	Chytridiomycota	
Rozella_allomyces_CSF55 EPZ32890.1	fungi	Cryptomycota	
Absidia_repens ORZ23163.1	fungi	Mucoromycota	
Bifiguratus_adelaidae OZJ02284.1	fungi	Mucoromycota	
Lichtheimia_corymbifera_JMRC_FSU_9682 CDH53810.1	fungi	Mucoromycota	
Mucor_ambiguus GAN00716.1	fungi	Mucoromycota	
Phycomyces_blakesleeanus_NRR1_1555_ XP_018290979.1	fungi	Mucoromycota	
Rhizopus_microsporus_ATCC_52813 XP_023471300.1	fungi	Mucoromycota	
Syncephalastrum_racemosum ORY99444.1	fungi	Mucoromycota	
Coemansia_reversa_NRR1_1564 PIA17930.1	fungi	Zoopagomycota	
Conidiobolus_coronatus_NRR1_28638 KXN70579.1	fungi	Zoopagomycota	
Linderina_pennisporea ORX67140.1	fungi	Zoopagomycota	
Pandora_neoaphidis APU66165.1	fungi	Zoopagomycota	

Table E-2: Extended table of all results and parameters tested. Prior and posterior divergence time estimates calculated under all model assumptions described in Table 4-1. Divergence date ranges are given in Ma.

Node	Root	Deep Fungi	Dikarya (AB Split)	Ascomy- cota	Basidiomy- cota	Bacteria	Gammapro- teobacteria	Vibionales	Betapro- teobacteria	Bac- teroidetes	Firmicutes	Deltapro- teobacteria	Actinobac- teria	Prior	Clock
Node #	1	3	5	6	70	147	148	165	155	174	181	186	215		
Calibrations															
Cal 1	1145-738		830-518												
Cal 2	1145-738		830-518	715-408	655-400										
Cal 3	1145-738		830-405												
Cal 4	1145-738		830-405	715-408	655-400										
Models															
1	814 (738-883)	804 (730-876)	760 (694-830)	651 (580-728)	674 (602-749)	695 (621-769)	668 (594-753)	427 (299-564)	567 (478-663)	577 (490-662)	544 (455-625)	447 (360-542)	509 (402-599)	U	LN
1p	817 (739-879)	809 (731-874)	764 (698-826)	680 (607-739)	686 (606-749)	726 (657-799)	691 (615-772)	443 (306-539)	583 (496-650)	605 (541-692)	584 (506-677)	548 (361-550)	543 (470-614)	U	LN
2	770 (738-819)	760 (722-806)	718 (677-762)	608 (552-659)	632 (597-655)	658 (616-710)	631 (580-684)	403 (284-511)	536 (470-614)	552 (487-622)	519 (450-588)	426 (350-501)	495 (415-567)	U	LN
2p	767 (738-807)	758 (725-801)	713 (680-749)	634 (601-679)	632 (603-653)	674 (637-715)	640 (594-687)	407 (305-510)	530 (466-590)	557 (500-618)	539 (475-595)	418 (359-499)	480 (415-538)	U	LN
3	817 (739-884)	808 (732-876)	763 (697-829)	656 (590-731)	677 (607-754)	698 (626-766)	670 (597-743)	426 (290-550)	566 (479-648)	584 (504-662)	550 (471-630)	452 (361-537)	514 (432-605)	U	LN
3p	818 (745-888)	811 (737-877)	766 (701-830)	683 (614-748)	692 (618-755)	726 (656-787)	695 (628-766)	477 (357-582)	586 (515-664)	601 (532-681)	573 (508-646)	439 (343-511)	536 (443-609)	U	LN
4	771 (738-813)	761 (722-803)	719 (682-762)	613 (561-659)	631 (599-655)	656 (612-708)	630 (581-682)	402 (289-514)	536 (458-600)	544 (476-614)	512 (441-579)	419 (345-498)	478 (408-550)	U	LN
4p	766 (738-804)	758 (726-797)	713 (682-751)	635 (598-677)	633 (606-655)	677 (642-719)	644 (598-689)	425 (305-525)	537 (473-596)	568 (512-631)	546 (489-606)	419 (339-507)	502 (432-573)	U	LN
5	800 (738-854)	793 (731-848)	775 (717-829)	559 (481-630)	628 (559-698)	618 (540-695)	516 (402-621)	192 (111-283)	339 (238-456)	447 (347-530)	416 (332-507)	278 (156-384)	372 (284-448)	U	UGAM
5p	845 (748-917)	836 (745-909)	763 (687-829)	590 (509-684)	640 (568-717)	708 (638-798)	592 (477-692)	281 (201-376)	394 (311-483)	526 (430-597)	496 (430-585)	355 (276-472)	471 (418-545)	U	UGAM
6	787 (738-837)	780 (729-833)	766 (720-828)	550 (480-618)	613 (566-655)	605 (537-672)	505 (393-605)	188 (113-278)	330 (223-442)	440 (359-525)	408 (323-498)	275 (174-389)	365 (290-435)	U	UGAM
6p	814 (742-909)	808 (730-902)	739 (676-824)	556 (477-621)	609 (544-653)	689 (619-792)	578 (486-655)	260 (183-369)	372 (307-472)	509 (429-596)	477 (395-546)	328 (254-415)	462 (383-527)	U	UGAM
7	800 (738-853)	794 (735-853)	775 (720-830)	560 (484-637)	626 (554-687)	616 (541-690)	511 (396-628)	189 (119-285)	328 (222-441)	446 (358-540)	413 (324-501)	276 (158-381)	370 (287-445)	U	UGAM
7p	811 (752-886)	806 (748-880)	755 (691-823)	577 (523-654)	634 (564-706)	689 (624-759)	575 (473-663)	257 (193-339)	377 (264-463)	512 (427-575)	488 (444-570)	328 (231-433)	452 (396-509)	U	UGAM
8	791 (738-845)	783 (732-842)	767 (716-825)	551 (479-622)	612 (563-655)	609 (542-680)	505 (393-604)	185 (102-266)	328 (218-430)	439 (350-524)	405 (329-494)	271 (163-381)	362 (291-443)	U	UGAM
8p	805 (750-865)	799 (737-854)	733 (672-790)	567 (514-642)	608 (556-647)	684 (631-740)	579 (484-649)	279 (202-379)	378 (279-469)	509 (450-564)	480 (419-534)	346 (267-438)	450 (393-507)	U	UGAM
9	1026 (891-1146)	926 (796-1048)	781 (700-830)	366 (267-457)	456 (349-572)	381 (290-479)	292 (200-390)	90 (57-125)	166 (110-219)	234 (172-294)	199 (149-250)	129 (80-177)	183 (138-236)	BD	UGAM
9p	1025 (889-1146)	911 (765-1034)	759 (655-829)	454 (359-565)	535 (433-630)	622 (490-734)	475 (365-612)	184 (135-247)	267 (198-359)	398 (310-489)	351 (268-426)	232 (163-299)	339 (270-420)	BD	UGAM
10	1071 (968-1146)	999 (902-1105)	798 (715-830)	428 (408-472)	494 (400-572)	424 (352-498)	333 (244-418)	98 (66-133)	191 (135-259)	273 (214-332)	232 (178-283)	148 (95-208)	210 (162-256)	BD	UGAM
10p	1024 (893-1143)	932 (812-1051)	770 (687-830)	441 (408-484)	550 (466-633)	637 (526-745)	494 (387-638)	192 (139-249)	284 (208-361)	409 (324-477)	367 (296-440)	246 (177-307)	347 (273-422)	BD	UGAM
11	1007 (872-1145)	910 (783-1031)	781 (686-830)	366 (276-447)	462 (369-570)	386 (292-482)	295 (206-404)	89 (59-121)	167 (115-220)	237 (174-302)	201 (156-252)	129 (80-178)	185 (141-237)	BD	UGAM
11p	1017 (864-1139)	918 (781-1061)	760 (656-829)	456 (345-561)	534 (422-627)	632 (496-737)	481 (357-607)	186 (132-249)	272 (190-344)	406 (318-522)	353 (282-430)	236 (167-314)	349 (272-427)	BD	UGAM
12	1062 (956-1146)	992 (890-1098)	798 (745-830)	429 (408-468)	495 (403-578)	424 (360-502)	331 (251-417)	98 (67-137)	189 (135-257)	274 (219-338)	232 (186-286)	149 (93-200)	209 (166-258)	BD	UGAM
12p	1065 (956-1145)	963 (843-1065)	772 (682-829)	443 (408-504)	544 (456-635)	648 (543-750)	400 (383-606)	189 (136-250)	279 (202-371)	409 (328-476)	369 (313-431)	242 (169-306)	347 (287-404)	BD	UGAM
13	990 (830-1141)	882 (747-1006)	769 (673-830)	494 (391-592)	550 (443-650)	544 (431-655)	499 (387-600)	187 (100-269)	361 (265-460)	349 (246-442)	322 (233-411)	247 (166-319)	290 (202-384)	BD	LN
13p	926 (776-1066)	868 (725-959)	765 (661-830)	580 (457-669)	653 (484-679)	582 (546-751)	246 (476-697)	413 (135-349)	433 (320-536)	291 (317-531)	360 (311-519)	291 (218-384)	360 (259-459)	BD	LN
14	1035 (890-1146)	871 (708-1004)	744 (633-830)	480 (409-561)	548 (441-648)	526 (428-619)	480 (381-581)	186 (99-280)	351 (258-444)	363 (285-440)	285 (265-400)	255 (194-326)	285 (213-356)	BD	LN
14p	959 (794-1117)	876 (747-990)	761 (661-829)	566 (470-659)	581 (495-655)	616 (526-713)	540 (444-643)	212 (120-302)	376 (287-466)	399 (315-484)	376 (286-449)	266 (185-332)	324 (257-400)	BD	LN
15	965 (796-1131)	860 (723-969)	760 (648-830)	510 (397-616)	552 (430-674)	509 (430-675)	204 (384-631)	378 (105-310)	366 (267-494)	335 (269-483)	258 (226-439)	304 (173-366)	275 (214-396)	BD	LN
15p	1003 (835-1146)	868 (704-1002)	749 (617-830)	548 (425-656)	572 (446-682)	614 (476-723)	542 (408-662)	219 (124-324)	380 (272-496)	389 (281-494)	366 (272-462)	259 (174-336)	317 (218-410)	BD	LN
16	1031 (879-1146)	880 (738-1001)	751 (649-830)	483 (408-563)	507 (440-655)	464 (412-612)	184 (365-565)	341 (97-273)	345 (249-439)	275 (200-429)	239 (240-385)	275 (172-304)	275 (203-353)	BD	LN
16p	925 (754-1049)	867 (736-980)	758 (660-829)	578 (484-654)	586 (498-655)	625 (530-722)	555 (439-647)	232 (138-342)	399 (301-506)	418 (317-513)	389 (301-473)	345 (200-344)	345 (268-433)	BD	LN

APPENDIX F

PHYLOGENETIC INVESTIGATIONS OF ACETOCLASTIC METHANOGENESIS

Acetoclastic methanogenesis is reported to account for two thirds of global microbial methane production (Ferry, 2010). This methane is produced by microbes in both fresh and marine aquatic systems. Because acetoclastic methanogenesis constitutes such a large fraction of the methane budget, the evolution of this pathway is an important event in the establishment of the modern carbon cycle. Phylogenetic analysis has shown that the two enzymes utilized by *Methanosarcina* to generate acetyl-CoA from acetate, acetate kinase (AckA) and phosphoacetyltransferase (Pta), were likely acquired from cellulolytic bacteria of the class *Clostridia* in a single HGT event (Figure 3-6) which had important implications for the carbon cycle during the end Permian (Rothman et al., 2014).

However, this event was dated using only the species tree of methanogens with highly conserved sequences, but without reconciling dates of events within the donor lineage. One approach to improving these dates is to evaluate HGT donor lineages in order to independently time this event with higher precision and to confirm that the transfer was related to the Permian-Triassic extinction. Also, this permits date calibrations from within *Clostridia* to be used, such as the origin of the cellulosome (Desvaux, 2005), a highly complex structure and one of the most efficient plant-degradation systems that must have arisen after lignocellulose appeared in the environment (Dassa et al., 2015). Fossil and molecular clock studies using a number of underlying assumptions indicate that vascular plants diverged in the Silurian, around 440 Ma (Heckman, 2001; Magallon et al., 2013; Sanderson et al., 2004). Nonetheless, independent age estimates for evolutionary events within the donor lineage *Clostridia* should more tightly constrain when methanogens acquired the ability to utilize acetate. Colonization of land plants and the evolution of lignin must predate the evolution of the cellulosome, providing a secondary time constraint.

Firmicutes also make up many of the bacterial taxa on the gene tree described in Chapter 3 of this thesis. More specifically, *Clostridia*, (*Clostridium botulinum* WP_075860993.1, *Clostridium carboxidivorans* WP_007060633.1, *Clostridium cellulovorans* WP_010076138.1, *Clostridium ljungdahlii* WP_063554956.1, *Clostridium purinilyticum* WP_050356052.1, *Clostridium scatologenes* WP_029160050.1, *Clostridium senegalense* WP_010292018.1, *Clostridium* sp KHNHs214 WP_035294784.1, *Clostridium straminisolvans* JCM 21531 WP_038286537.1, *Clostridium tyrobutyricum* WP_017751388.1) are present. This representation is interesting from the perspective of understanding the evolutionary history of methanogenic pathways, as another pathway of methanogenesis, the acetoclastic pathway, has been implicated in the HGT from cellulolytic *Clostridia* to acetoclastic methanogens (Fournier and Gogarten, 2008). Figure F-1 illustrates the hypothesized relationship between cellulolytic *Clostridia* and acetoclastic methanogens. There may be considerable overlap between the Firmicutes in both acetoclastic and methylophilic methanogenesis.

While acetoclastic methanogenesis is more widely distributed in sampled environments cultured microbes, the role of methylophilic methanogenesis is becoming increasingly ap-

preciated, particularly in marine environments. Figure F-2 depicts a gene tree generated from AckA and Pta proteins of methanogens and their closest relatives. This tree expands upon previous work and is consistent with the hypothesis that acetoclastic methanogenesis evolved via HGT (Rothman et al., 2014). This tree also illustrates overlap between donor lineages of AckA and Pta proteins and those of mtaC proteins. In the environment, methanol is produced by the degradation of plant cell walls (Sousa et al., 2018). Thus, it stands to reason that there may be overlap between organisms that can degrade plant components and those that use methanol; they both occupy similar environments and overlapping substrates. Indeed, cellulolytic Bacteria such as *Pseudobacteroides cellulosolvens* have both a methyltransferase protein (WP_036939029.1) (Dassa et al., 2015) and AckA Pta genes. Thus, timing the origin of cellulolytic *Clostridia* may serve the dual role of aiding in understanding the origin of two environmentally-important methanogenic pathways.

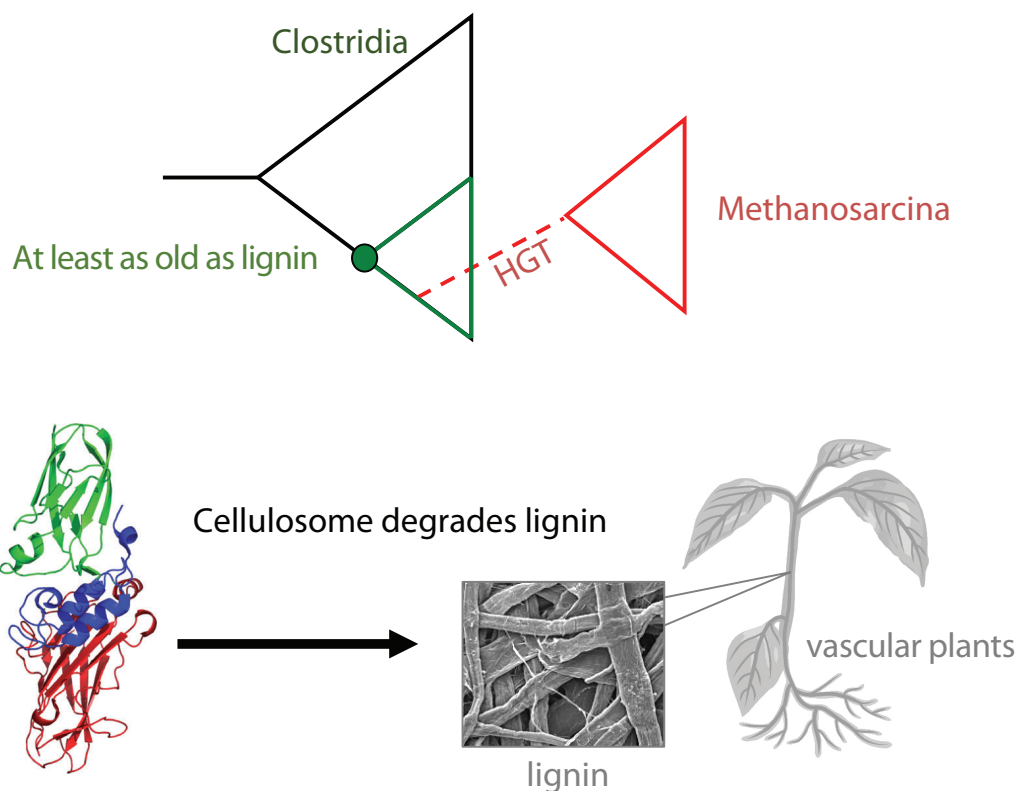


Figure F-1: The relationship between donor and recipient lineages of acetoclastic methanogens. Acetoclastic methanogenesis is present only in archaea and evolved in *Methanosarcina* from a clade of cellulolytic bacteria. The donor lineage of acetoclastic methanogenesis possessed a cellulosome. The cellulosome is a plant cell wall degrading apparatus in anaerobic bacteria. The cellulosome is a complex structure (thereby containing a lot of information), moreover, because it takes so much maintenance energy to preserve a structure as big as this, we assume that it must have evolved after lignocellulose appeared in the environment. The evolution of this structure itself is a fundamental biological process that is critical to the cycling of carbon between microbes, herbivores, and plants. If the HGT to *Methanosarcina* is from within a clade of cellulolytic bacteria (green), it can be inferred that the donor lineage is younger than the origin of cellulose, and therefore the recipient lineage is also younger than the origin of cellulose. This is useful, because fossils and molecular clock studies indicate that the source of lignocellulose, vascular plants diverged in the Silurian, around 440 Ma.

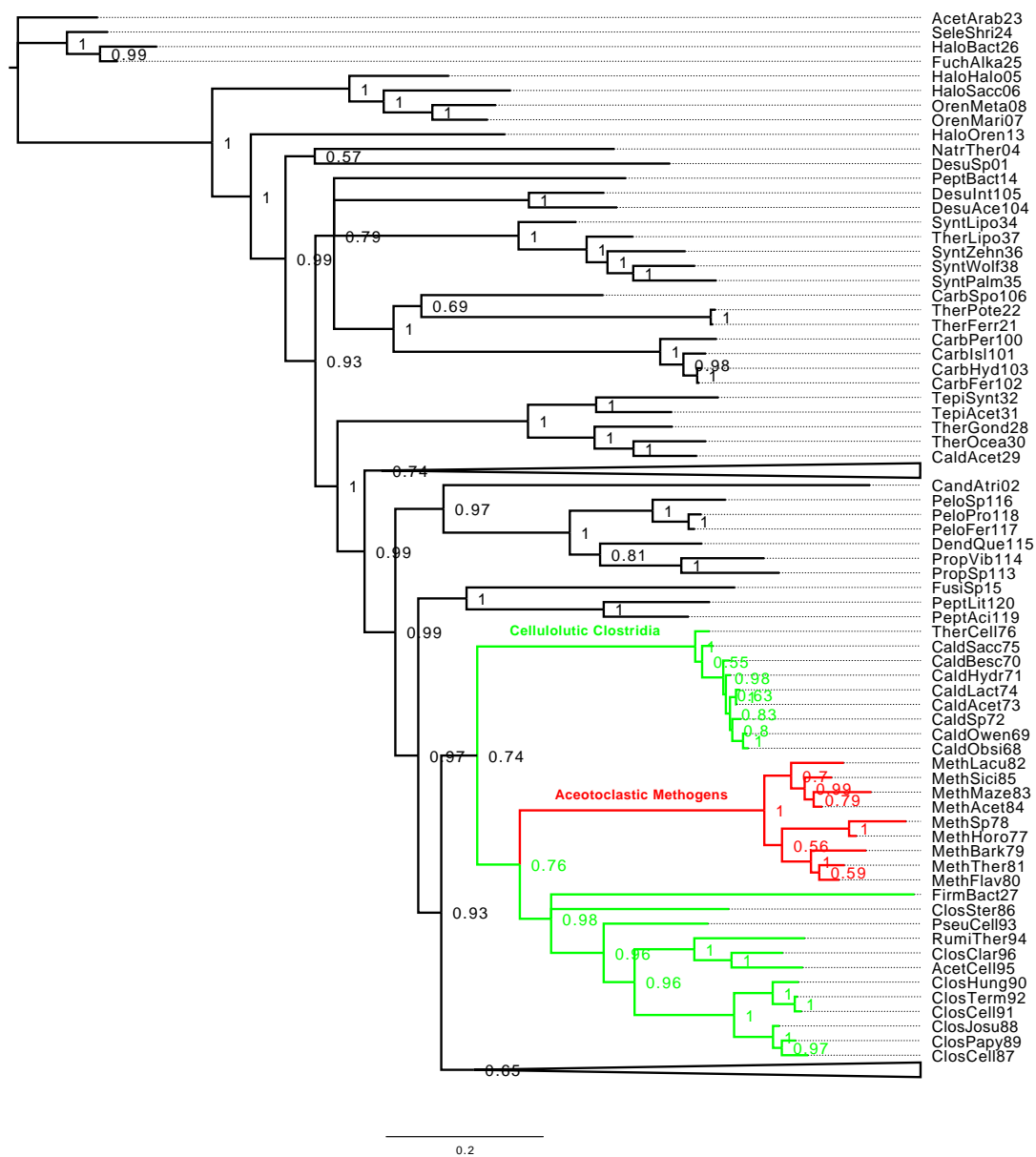


Figure F-2: Concatenated AckA Pta gene tree of methanogens and their closest relatives.

APPENDIX G

REFERENCES

BIBLIOGRAPHY

- Affek, H. P. and J. M. Eiler, 2006. Abundance of mass 47 CO₂ in urban air, car exhaust, and human breath. *Geochimica et Cosmochimica Acta* 70(1): 1–12. doi:10.1016/j.gca.2005.08.021.
- Affek, H. P., X. Xu and J. M. Eiler, 2007. Seasonal and diurnal variations of ¹³C¹⁸O¹⁶O in air: Initial observations from Pasadena, CA. *Geochimica et Cosmochimica Acta* 71(21): 5033–5043. doi:10.1016/j.gca.2007.08.014.
- Agalidis, I., A. Ivancich, T. A. Mattioli and F. Reiss-Husson, 1997. Characterization of the *Rhodocyclus tenuis* photosynthetic reaction center. *Biochimica et Biophysica Acta* 31: 31–46.
- Alcolombri, U., S. Ben-Dor, E. Feldmesser, Y. Levin, D. S. Tawfik and A. Vardi, 2015. Identification of the algal dimethyl sulfide-releasing enzyme: A missing link in the marine sulfur cycle. *Science* 348(6242): 1466–1469. doi:10.1126/science.aab1586.
- Alperin, M. and N. Blair, 1992. Factors that control the stable carbon isotopic composition of methane produced in an anoxic marine sediment. *Global Biogeochemical Cycles* 6(3): 271–291.
- Amin, M., R. C. Price and W. H. Saunders, 1988. Isotope effects on isotope effects. failure of the rule of the geometric mean as evidence for tunneling. *Journal of the American Chemical Society* 110(12): 4085–4086.
- Andam, C. P., D. Williams and J. P. Gogarten, 2010. Natural taxonomy in light of horizontal gene transfer. *Biology and Philosophy* 25(4): 589–602. doi:10.1007/s10539-010-9212-8.
- Anderson, I. J., H. Sun, A. Lapidus, A. Copeland, T. G. Del Rio, H. Tice, E. Dalin, S. Lucas, K. Barry, M. Land, P. Richardson, H. Huber and N. C. Kyrpides, 2009. Complete genome sequence of *staphylothermus marinus* stetter and fiala 1986 type strain F1. *Standards in Genomic Sciences* 1(2): 183–188. doi:10.4056/sigs.30527.
- Ayers, G. P. and J. M. Caaney, 2007. The CLAW hypothesis: a review of the major developments. *Environmental Chemistry* 4(6): 366. doi:10.1071/EN07080.
- Bai, Y., V. G. H. Eijsink, A. M. Kielak, J. A. van Veen and W. de Boer, 2016. Genomic comparison of chitinolytic enzyme systems from terrestrial and aquatic bacteria. *Environmental Microbiology* 18(1): 38–49. doi:10.1111/1462-2920.12545.
- Baker, B. J., B. J. Baker, J. H. Saw, A. E. Lind, C. S. Lazar, K.-u. Hinrichs, A. P. Teske and T. J. G. Ettema, 2016. Genomic inference of the metabolism of cosmopolitan subsurface archaea, Hadesarchaea. *Nature Microbiology* (February): 1–7. doi:10.1038/nmicrobiol.2016.2.

- Balabane, M., E. Galimov, M. Hermann and R. Letolle, 1987. Hydrogen and carbon isotope fractionation during experimental production of bacterial methane. *Organic Geochemistry* 11(2): 115–119.
- Balch, W., G. Fox, L. Magrum, C. Woese and R. Wolfe, 1979. Methanogens: Reevaluation of a Unique Biological Group. *Microbiology Reviews* 43(2): 260–296.
- Baldassare, F. J., M. A. McCaffrey and J. A. Harper, 2014. A geochemical context for stray gas investigations in the northern appalachian basin: Implications of analyses of natural gases from neogene-through devonian-age strata. *AAPG bulletin* 98(2): 341–372.
- Baptiste, E., C. Brochier and Y. Boucher, 2005. Higher-level classification of the Archaea: evolution of methanogenesis and methanogens. *Archaea (Vancouver, B.C.)* 1(5): 353–363.
- Bar-On, Y. M., R. Phillips, R. Milo and P. G. Falkowski, 2018. The biomass distribution on Earth doi:10.1073/pnas.1711842115.
- Bates, B. L., J. C. McIntosh, K. A. Lohse and P. D. Brooks, 2011. Influence of groundwater flowpaths, residence times and nutrients on the extent of microbial methanogenesis in coal beds: Powder river basin, usa. *Chemical Geology* 284(1-2): 45–61.
- Battistuzzi, F. U., A. Feijao and S. B. Hedges, 2004. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC evolutionary biology* 4(1): 44. doi:10.1186/1471-2148-4-44.
- Bayer, B., J. Vojvoda, P. Offre, R. J. Alves, N. H. Elisabeth, J. A. Garcia, J. M. Volland, A. Srivastava, C. Schleper and G. J. Herndl, 2016. Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME Journal* 10(5): 1051–1063. doi:10.1038/ismej.2015.200.
- Berbee, M. L., T. Y. James and C. Strullu-Derrien, 2017. Early Diverging Fungi: Diversity and Impact at the Dawn of Terrestrial Life. *Annual Review of Microbiology* 71(1): 41–60. doi:10.1146/annurev-micro-030117-020324.
- Berbee, M. L. and J. W. Taylor, 2010. Dating the molecular clock in fungi – how close are we? *Fungal Biology Reviews* 24(1-2): 1–16. doi:10.1016/j.fbr.2010.03.001.
- Berlemont, R. and A. C. Martiny, 2015. Genomic potential for polysaccharide deconstruction in bacteria. *Applied and Environmental Microbiology* 81(4): 1513–1519. doi:10.1128/AEM.03718-14.
- Bigeleisen, J., 1949. The relative reaction velocities of isotopic molecules. *The journal of chemical physics* 17(8): 675–678.
- Bigeleisen, J., 1955. Statistical mechanics of isotopic systems with small quantum corrections. i. general considerations and the rule of the geometric mean. *The Journal of Chemical Physics* 23(12): 2264–2267.
- Blair, N. E. and W. D. Carter, 1992. The carbon isotope biogeochemistry of acetate from a methanogenic marine sediment. *Geochimica et Cosmochimica Acta* 56(3): 1247–1258. doi:10.1016/0016-7037(92)90060-V.

- Blaser, M. and R. Conrad, 2016. Stable carbon isotope fractionation as tracer of carbon cycling in anoxic soil ecosystems. *Current Opinion in Biotechnology* 41: 122–129. doi:10.1016/j.copbio.2016.07.001.
- Bo, M., G. Bavestrello, D. Kurek, S. Paasch, E. Brunner, R. Born, R. Galli, A. L. Stelling, V. N. Sivkov, O. V. Petrova, D. Vyalikh, K. Kummer, S. L. Molodtsov, D. Nowak, J. Nowak and H. Ehrlich, 2012. Isolation and identification of chitin in the black coral *Parantipathes larix* (Anthozoa: Cnidaria). *International Journal of Biological Macromolecules* 51(1-2): 129–137. doi:10.1016/j.ijbiomac.2012.04.016.
- Boetius, A. and E. Suess, 2004. Hydrate Ridge: a natural laboratory for the study of microbial life fueled by methane from near-surface gas hydrates. *Chemical Geology* 205(3-4): 291–310. doi:10.1016/j.chemgeo.2003.12.034.
- Bontognali, T. R. R., A. L. Sessions, A. C. Allwood, W. W. Fischer, J. P. Grotzinger, R. E. Summons and J. M. Eiler, 2012. Sulfur isotopes of organic matter preserved in 3.45-billion-year-old stromatolites reveal microbial metabolism. *Proceedings of the National Academy of Sciences of the United States of America* 109(38): 15146–51. doi:10.1073/pnas.1207491109.
- Borrel, G., N. Parisot, H. M. Harris, E. Peyretilade, N. Gaci, W. Tottey, O. Bardot, K. Raymann, S. Gribaldo, P. Peyret, P. W. O’Toole and J.-F. Brugère, 2014. Comparative genomics highlights the unique biology of *Methanomassiliicoccales*, a *Thermoplasmatales*-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC Genomics* 15(1): 679. doi:10.1186/1471-2164-15-679.
- Braakman, R., M. J. Follows and S. W. Chisholm, 2017. Metabolic evolution and the self-organization of ecosystems. *Proceedings of the National Academy of Sciences* 114(15): E3091–E3100. doi:10.1073/pnas.1619573114.
- Brocks, J. J. and A. Pearson, 2005. Building the Biomarker Tree of Life. *Reviews in Mineralogy & Geochemistry* 59: 223–258.
- Bromham, L., S. Duchêne, X. Hua, A. M. Ritchie, D. A. Duchêne and S. Y. W. Ho, 2018. Bayesian molecular dating: opening up the black box. *Biological Reviews* 93(2): 1165–1191. doi:10.1111/brv.12390.
- Burke, R. A., 1993. Possible influence of hydrogen concentration on microbial methane stable hydrogen isotopic composition. *Chemosphere* 26(1-4): 55–67. doi:10.1016/0045-6535(93)90412-X.
- Burke, S. A., S. L. Lo and J. A. Krzycki, 1998. Clustered genes encoding the methyl-transferases of methanogenesis from monomethylamine. *Journal of bacteriology* 180(13): 3432–40.
- Burke Jr, R. A., C. S. Martens and W. M. Sackett, 1988. Seasonal variations of d/h and $^{13}\text{C}/^{12}\text{C}$ ratios of microbial methane in surface sediments. *Nature* 332(6167): 829.
- Burruss, R. and C. Laughrey, 2010. Carbon and hydrogen isotopic reversals in deep basin gas: Evidence for limits to the stability of hydrocarbons. *Organic Geochemistry* 41(12): 1285–1296.

- Cardace, D., T. Hoehler, T. McCollom, M. Schrenk, D. Carnevale, M. Kubo and K. Twing, 2013. Establishment of the coast range ophiolite microbial observatory (cromo): drilling objectives and preliminary outcomes. *Scientific Drilling* 16: 45–55. doi:10.5194/sd-16-45-2013.
- Caspi, R., T. Altman, R. Billington, K. Dreher, H. Foerster, C. A. Fulcher, T. A. Holland, I. M. Keseler, A. Kothari, A. Kubo, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D. S. Weaver, D. Weerasinghe, P. Zhang and P. D. Karp, 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research* 42(D1): 459–471. doi:10.1093/nar/gkt1103.
- Castelle, C. J. and J. F. Banfield, 2018. Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* 172(6): 1181–1197. doi:10.1016/j.cell.2018.02.016.
- Cauchie, H.-M., 2002. Chitin production by arthropods in the hydrosphere. *Hydrobiologia* 470(1/3): 63–95. doi:10.1023/A:1015615819301.
- Cha, Y., C. Murray and J. Klinman, 1989. Hydrogen tunneling in enzyme reactions. *Science* 243(4896): 1325–1330. doi:10.1126/science.2646716.
- Chang, T. C. and I. Stergiopoulos, 2015. Evolutionary analysis of the global landscape of protein domain types and domain architectures associated with family 14 carbohydrate-binding modules. *FEBS Letters* 589(15): 1813–1818. doi:10.1016/j.febslet.2015.05.048.
- Charlson, R. J., J. E. Lovelock, M. O. Andreae and S. G. Warren, 1987. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* 326: 655–661. doi:10.1038/326655a0.
- Chen, S. L., M. R. A. Blomberg and P. E. M. Siegbahn, 2012. How is methane formed and oxidized reversibly when catalyzed by Ni-containing methyl-coenzyme M reductase? *Chemistry - A European Journal* 18(20): 6309–6315. doi:10.1002/chem.201200274.
- Cheng, W.-H., 1994. *Methanol production and use*. CRC Press.
- Colgate, S. A., 1979. Supernovae as a standard candle for cosmology. *The Astrophysical Journal* 232: 404. doi:10.1086/157300.
- Conrad, R., 2005. Quantification of methanogenic pathways using stable carbon isotopic signatures: a review and a proposal. *Organic Geochemistry* 36(5): 739–752. doi:10.1016/j.orggeochem.2004.09.006.
- Conrad, R., P. Claus and P. Casper, 2009. Characterization of stable isotope fractionation during methane production in the sediment of a eutrophic lake, Lake Dagow, Germany 54(2): 457–471.
- Coplen, T. B., 2011. Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results. *Rapid Communications in Mass Spectrometry* 25(17): 2538–2560. doi:10.1002/rcm.5129.

- Curson, A. R. J., J. D. Todd, M. J. Sullivan and A. W. B. Johnston, 2011. Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nature Reviews Microbiology* 9(12): 849–859. doi:10.1038/nrmicro2653.
- Dacey, J. W. H. and S. G. Wakeham, 1986. Oceanic dimethyl sulfide: production during zooplankton grazing on phytoplankton. *Science* 233(13): 1314–1316.
- Daley, A. C., J. B. Antcliffe, H. B. Drage and S. Pates, 2018. Early fossil record of Euarthropoda and the Cambrian Explosion pages 1–9. doi:10.1073/pnas.1719962115.
- Daniels, L., G. Fulton, R. Spencer and W. Orme-Johnson, 1980. Origin of hydrogen in methane produced by methanobacterium thermoautotrophicum. *Journal of Bacteriology* 141(2): 694–698.
- Darriba, D., G. L. Taboada, R. Doallo and D. Posada, 2017. Europe PMC Funders Group ProtTest 3 : fast selection of best-fit models of protein evolution. *Bioinformatics* 27(8): 1164–1165. doi:10.1093/bioinformatics/btr088.ProtTest.
- Dassa, B., S. Utturkar, R. A. Hurt, D. M. Klingeman, M. Keller, J. Xu, Y. H. K. Reddy, I. Borovok, I. Rozman Grinberg, R. Lamed, O. Zhivin, E. A. Bayer and S. D. Brown, 2015. Near-Complete Genome Sequence of the Cellulolytic Bacterium Bacteroides (Pseudobacteroides) cellulosolvens ATCC 35603. *Genome Announcements* 3(5): e01022–15. doi:10.1128/genomeA.01022-15.
- David, L. A. and E. J. Alm, 2011. Rapid evolutionary innovation during an Archaeal genetic expansion. *Nature* 469(7328): 93–96. doi:10.1038/nature09649.
- Davin, A. A., E. Tannier, T. A. Williams, B. Boussau, V. Daubin and G. J. Szollosi, 2017. Gene transfers, like fossils, can date the Tree of Life. *Doi.Org* page 193813. doi:10.1101/193813.
- De Graaf, W., P. Wellsbury, R. J. Parkes and T. E. Cappenberg, 1996. Comparison of acetate turnover in methanogenic and sulfate-reducing sediments by radiolabeling and stable isotope labeling and by use of specific inhibitors: evidence for isotopic exchange. *Applied and Environmental Microbiology* 62(3): 772–777.
- Defliese, W. F. and K. C. Lohmann, 2015. Non-linear mixing effects on mass-47 CO₂ clumped isotope thermometry: Patterns and implications. *Rapid communications in mass spectrometry : RCM* 29(9): 901–9. doi:10.1002/rcm.7175.
- Deppenmeier, U., A. Johann, T. Hartsch, R. Merkl, R. A. Schmitz, R. Martinez-Arias, A. Henne, A. Wiezer, S. Bäumer, C. Jacobi, H. Brüggemann, T. Lienard, A. Christmann, M. Bömeke, S. Steckel, A. Bhattacharyya, A. Lykidis, R. Overbeek, H.-P. Klenk, R. P. Gunsalus, H.-J. Fritz and G. Gottschalk, 2002. The Genome of Methanosarcina mazei: Evidence for Lateral Gene Transfer Between Bacteria and Archaea JMMB Research Article. *J. Mol. Microbiol. Biotechnol* 4(4): 453–461.
- Desvaux, M., 2005. Clostridium cellulolyticum : model organism of mesophilic cellulolytic clostridia. *FEMS Microbiology Reviews* 29(4): 741–764. doi:10.1016/j.femsre.2004.11.003.
- Dickson, D. M. J. and G. O. Kirst, 1986. The role of B-dimethylsulphoniopropionate, glycine betaine and homarine in the osmoacclimation of Platymonas subcordiformis. *Planta* 167(4): 536–543. doi:10.1007/BF00391230.

- Dong, M., T. D. Gonzalez, M. M. Klems, L. M. Steinberg, W. Chen, E. T. Papoutsakis and B. J. Bahnson, 2017. In vitro methanol production from methyl coenzyme M using the *Methanosarcina barkeri* MtaABC protein complex. *Biotechnology Progress* 33(5): 1243–1249. doi:10.1002/btpr.2503.
- Donoghue, P. C. J. and M. J. Benton, 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology and Evolution* 22(8): 424–431. doi:10.1016/j.tree.2007.05.005.
- Donoghue, P. C. J. and Z. Yang, 2016. The evolution of methods for establishing evolutionary timescales. *Philosophical Transactions of the Royal Society of London B* 371: 20160020. doi:10.1098/rstb.2016.0020.
- Dos Reis, M., 2018. Fossil-free dating. *Nature Ecology and Evolution* 2(May): 1–2. doi:10.1038/s41559-018-0532-4.
- dos Reis, M., P. C. J. Donoghue and Z. Yang, 2016. Bayesian molecular clock dating of species divergences in the genomics era. *Nature Reviews Genetics* 17(2): 71–80. doi:10.1038/nrg.2015.8.
- Dos Reis, M., Y. Thawornwattana, K. Angelis, M. J. Telford, P. C. Donoghue and Z. Yang, 2015. Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. *Current Biology* 25(22): 2939–2950. doi:10.1016/j.cub.2015.09.066.
- Douglas, P., D. Stolper, D. Smith, K. Walter Anthony, C. Paull, S. Dallimore, M. Wik, P. Crill, M. Winterdahl, J. Eiler and A. Sessions, 2016. Diverse origins of Arctic and Sub-arctic methane point source emissions identified with multiply-substituted isotopologues. *Geochimica et Cosmochimica Acta* 188: 163–188. doi:10.1016/j.gca.2016.05.031.
- Douglas, P. M., D. A. Stolper, J. M. Eiler, A. L. Sessions, M. Lawson, Y. Shuai, A. Bishop, O. G. Podlaha, A. A. Ferreira, E. V. Santos Neto, M. Niemann, A. S. Steen, L. Huang, L. Chimiak, D. L. Valentine, J. Fiebig, A. J. Luhmann, W. E. Seyfried, G. Etiope, M. Schoell, W. P. Inskeep, J. J. Moran and N. Kitchen, 2017. Methane clumped isotopes: Progress and potential for a new isotopic tracer. *Organic Geochemistry* 113: 262–282. doi:10.1016/j.orggeochem.2017.07.016.
- Dridi, B., M. L. Fardeau, B. Ollivier, D. Raoult and M. Drancourt, 2012. *Methanomasiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology* 62(8): 1902–1907. doi:10.1099/ijs.0.033712-0.
- Drummond, A. J., S. Y. W. Ho, M. J. Phillips and A. Rambaut, 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4(5): 699–710. doi:10.1371/journal.pbio.0040088.
- Dupont, C. L., A. Butcher, R. E. Valas, P. E. Bourne and G. Caetano-Anollés, 2010. History of biological metal utilization inferred through phylogenomic analysis of protein structures. *Proceedings of the National Academy of Sciences of the United States of America* 107(23): 10567–10572. doi:10.1073/pnas.0912491107.
- Dupont, C. L., S. Yang, B. Palenik and P. E. Bourne, 2006. Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proceedings of the National Academy of Sciences* 103(47): 17822–17827. doi:10.1073.

- Eby, R. and T. Singleton, 1983. Chapter 10: Methanol Carbonylation to Acetic Acid. In B. E. Leach, editor, *Applied Industrial Catalysis*, page 275. Academic Press. doi: <https://doi.org/10.1016/B978-0-12-440201-0.50003-5>.
- Edgar, R. C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5): 1792–1797. doi:10.1093/nar/gkh340.
- Edwards, S. V., 2009. Is a new and general theory of molecular systematics emerging? *Evolution* 63(1): 1–19. doi:10.1111/j.1558-5646.2008.00549.x.
- Eiler, J. M., B. Bergquist, I. Bourg, P. Cartigny, J. Farquhar, A. Gagnon, W. Guo, I. Halevy, A. Hofmann, T. E. Larson, N. Levin, E. a. Schauble and D. Stolper, 2014. Frontiers of Stable Isotope Geoscience. *Chemical Geology* doi:10.1016/j.chemgeo.2014.02.006.
- Eiler, J. M. and E. Schauble, 2004. $^{18}\text{O}/^{13}\text{C}/^{16}\text{O}$ in Earth’s atmosphere. *Geochimica et Cosmochimica Acta* 68(23): 4767–4777. doi:10.1016/j.gca.2004.05.035.
- Elkins, J. G., M. Podar, D. E. Graham, K. S. Makarova, Y. Wolf, L. Randau, B. P. Hedlund, C. Brochier-Armanet, V. Kunin, I. Anderson, A. Lapidus, E. Goltsman, K. Barry, E. V. Koonin, P. Hugenholtz, N. Kyrpides, G. Wanner, P. Richardson, M. Keller and K. O. Stetter, 2008. A korarchaeal genome reveals insights into the evolution of the Archaea. *Proceedings of the National Academy of Sciences of the United States of America* 105(23): 8102–8107. doi:10.1073/pnas.0801980105.
- Emerling, C. A., F. Delsuc and M. W. Nachman, 2018. Chitinase genes (CHIAs) provide genomic footprints of a post-Cretaceous dietary radiation in placental mammals. *Science Advances* 4(5): 1–10. doi:10.1126/sciadv.aar6478.
- Ermler, U., W. Grabarse, S. Shima, M. Goubeaud and R. K. Thauer, 1997. Crystal structure of methyl coenzyme M reductase: The key enzyme of biological methane formation. *Science* 278(5342): 1457–1462. doi:10.1126/science.278.5342.1457.
- Etiope, G. and B. Sherwood Lollar, 2013. Abiotic methane on earth. *Reviews of Geophysics* 51(2): 276–299.
- Evans, P. N., D. H. Parks, G. L. Chadwick, S. J. Robbins, V. J. Orphan, S. D. Golding and G. W. Tyson, 2015. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* 350(6259): 434–438. doi:10.1126/science.aac7745.
- Falkowski, P. G., T. Fenchel and E. F. Delong, 2008. The microbial engines that drive Earth’s biogeochemical cycles. *Science (New York, N.Y.)* 320(5879): 1034–9. doi:10.1126/science.1153213.
- Farag, I. F., J. P. Davis, N. H. Youssef and M. S. Elshahed, 2014. Global patterns of abundance, diversity and community structure of the aminicenantes (Candidate Phylum OP8). *PLoS ONE* 9(3). doi:10.1371/journal.pone.0092139.
- Farquhar, J., D. T. Johnston and B. A. Wing, 2007. Implications of conservation of mass effects on mass-dependent isotope fractionations: Influence of network structure on sulfur isotope phase space of dissimilatory sulfate reduction. *Geochimica et Cosmochimica Acta* 71(24): 5862–5875. doi:10.1016/j.gca.2007.08.028.

- Ferguson, D. J., N. Gorlatova, D. A. Grahame and J. A. Krzycki, 2000. Reconstitution of dimethylamine:coenzyme M methyl transfer with a discrete corrinoid protein and two methyltransferases purified from *Methanosarcina barkeri*. *Journal of Biological Chemistry* 275(37): 29053–29060. doi:10.1074/jbc.M910218199.
- Ferguson, D. J. and J. A. Krzycki, 1997. Reconstitution of trimethylamine-dependent coenzyme M methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase II from *Methanosarcina barkeri*. *Journal of Bacteriology* 179(3): 846–852. doi:10.1128/jb.179.3.846-852.1997.
- Ferguson, T., J. A. Soares, T. Lienard, G. Gottschalk and J. A. Krzycki, 2009. RamA, a protein required for reductive activation of corrinoid-dependent methylamine methyltransferase reactions in methanogenic archaea. *Journal of Biological Chemistry* 284(4): 2285–2295. doi:10.1074/jbc.M807392200.
- Ferry, J. G., 2010. How to make a living by exhaling methane. *Annual review of microbiology* 64: 453–473. doi:10.1146/annurev.micro.112408.134051.
- Fischer, D., J. M. Mogollón, M. Strasser, T. Pape, G. Bohrmann, N. Fekete, V. Spiess and S. Kasten, 2013. Subduction zone earthquake as potential trigger of submarine hydrocarbon seepage. *Nature Geoscience* 6(8): 647–651. doi:10.1038/ngeo1886.
- Flores, R. M., C. A. Rice, G. D. Stricker, A. Warden and M. S. Ellis, 2008. Methanogenic pathways of coal-bed gas in the powder river basin, united states: the geologic factor. *International Journal of Coal Geology* 76(1-2): 52–75.
- Floudas, D., M. Binder, R. Riley, K. Barry, R. A. Blanchette, B. Henrissat, A. T. Martinez, R. Otillar, J. W. Spatafora, J. S. Yadav, A. Aerts, I. Benoit, A. Boyd, A. Carlson, A. Copeland, P. M. Coutinho, R. P. de Vries, P. Ferreira, K. Findley, B. Foster, J. Gaskell, D. Glotzer, P. Gorecki, J. Heitman, C. Hesse, C. Hori, K. Igarashi, J. A. Jurgens, N. Kallen, P. Kersten, A. Kohler, U. Kues, T. K. A. Kumar, A. Kuo, K. LaButti, L. F. Larrondo, E. Lindquist, A. Ling, V. Lombard, S. Lucas, T. Lundell, R. Martin, D. J. McLaughlin, I. Morgenstern, E. Morin, C. Murat, L. G. Nagy, M. Nolan, R. A. Ohm, A. Patyshakuliyeva, A. Rokas, F. J. Ruiz-Duenas, G. Sabat, A. Salamov, M. Samejima, J. Schmutz, J. C. Slot, F. St. John, J. Stenlid, H. Sun, S. Sun, K. Syed, A. Tsang, A. Wiebenga, D. Young, A. Pisabarro, D. C. Eastwood, F. Martin, D. Cullen, I. V. Grigoriev and D. S. Hibbett, 2012. The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science* 336(6089): 1715–1719. doi:10.1126/science.1221748.
- Fournier, G. P., C. P. Andam and J. P. Gogarten, 2015. Ancient horizontal gene transfer and the last common ancestors. *BMC Evolutionary Biology* 15(1): 70. doi:10.1186/s12862-015-0350-0.
- Fournier, G. P. and J. P. Gogarten, 2008. Evolution of Acetoclastic Methanogenesis in *Methanosarcina* via Horizontal Gene Transfer from Cellulolytic Clostridia. *Journal of Bacteriology* 190(3): 1124–1127. doi:10.1128/JB.01382-07.
- Fournier, G. P., J. Huang and J. P. Gogarten, 2009. Horizontal gene transfer from extinct and extant lineages: biological innovation and the coral of life. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364(1527): 2229–39. doi:10.1098/rstb.2009.0033.

- Fricke, W. F., H. Seedorf, A. Henne, M. Kru, H. Liesegang, R. Hedderich, G. Gottschalk and R. K. Thauer, 2006. The Genome Sequence of *Methanosphaera stadtmanae* Reveals Why This Human Intestinal Archaeon Is Restricted to Methanol and H₂ for Methane Formation and ATP Synthesis 188(2): 642–658. doi:10.1128/JB.188.2.642.
- Funkhouser, J. D. and N. N. Aronson, 2007. Chitinase family GH18: Evolutionary insights from the genomic history of a diverse protein family. *BMC Evolutionary Biology* 7: 1–16. doi:10.1186/1471-2148-7-96.
- Galagan, J. E., C. Nusbaum, A. Roy, M. G. Endrizzi, P. Macdonald, W. FitzHugh, S. Calvo, R. Engels, S. Smirnov, D. Atnoor, A. Brown, N. Allen, J. Naylor, N. Stange-Thomann, K. DeArellano, R. Johnson, L. Linton, P. McEwan, K. McKernan, J. Talamas, A. Tirrell, W. Ye, A. Zimmer, R. D. Barber, I. Cann, D. E. Graham, D. A. Grahame, A. M. Guss, R. Hedderich, C. Ingram-Smith, H. C. Kuettner, J. A. Krzycki, J. A. Leigh, W. Li, J. Liu, B. Mukhopadhyay, J. N. Reeve, K. Smith, T. A. Springer, L. A. Umayam, O. White, R. H. White, E. Conway de Macario, J. G. Ferry, K. F. Jarrell, H. Jing, A. J. L. Macario, I. Paulsen, M. Pritchett, K. R. Sowers, R. V. Swanson, S. H. Zinder, E. Lander, W. W. Metcalf and B. Birren, 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome research* 12(4): 532–42. doi:10.1101/gr.223902.
- Glass, J. B. and V. J. Orphan, 2012. Trace Metal Requirements for Microbial Enzymes Involved in the Production and Consumption of Methane and Nitrous Oxide. *Frontiers in Microbiology* 3(February): 1–20. doi:10.3389/fmicb.2012.00061.
- Glenner, H., P. F. Thomsen, M. B. Hebsgaard, M. V. Sørensen and E. Willerslev, 2006. The Origin of Insects. *Science* 314(5807): 1883–1884. doi:10.1126/science.1129844.
- Gogarten, J. P., 1995. The early evolution of cellular life. *Trends in ecology & evolution* 10(4): 147–51. doi:10.1016/S0169-5347(00)89024-2.
- Gogarten, J. P. and J. P. Townsend, 2005. Horizontal gene transfer, genome innovation and evolution. *Nature reviews. Microbiology* 3(9): 679–687. doi:10.1038/nrmicro1204.
- Gold, D. A., A. Caron, G. P. Fournier and R. E. Summons, 2017. Paleoproterozoic sterol biosynthesis and the rise of oxygen. *Nature* doi:10.1038/nature21412.
- Grabarse, W., F. Mahlert, S. Shima, R. K. Thauer and U. Ermler, 2000. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: Unusual amino acid modification, conservation and adaptation. *Journal of Molecular Biology* 303(2): 329–344. doi:10.1006/jmbi.2000.4136.
- Graber, J. R. and J. a. Breznak, 2004. Physiology and Nutrition of *Treponema primitia*, an H₂/ CO₂-Acetogenic Spirochete from Termite Hindguts. *Applied and Environmental Microbiology* 70(3): 1307–1314. doi:10.1128/AEM.70.3.1307-1314.2004.
- Green, T. K. and A. D. Hatton, 2014. The Claw Hypothesis: A New Perspective on the Role of Biogenic Sulphur in the Regulation of Global Climate. In *Oceanography and Marine Biology*, volume 52, pages 315–336. doi:10.1201/b17143-7.
- Gribaldo, S. and C. Brochier-Armanet, 2006. The origin and evolution of Archaea: a state of the art. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 361(1470): 1007–1022. doi:10.1098/rstb.2006.1841.

- Gruen, D., D. Wang, L. Stewart, J. Holden and S. Ono, 2014. Clumped isotope systematics of biogenic methane. *Mineralogical Magazine* 77(5): 2014.
- Gruen, D. S., D. T. Wang, M. Könneke, B. D. Topçuoğlu, L. C. Stewart, T. Goldhammer, J. F. Holden, K.-U. Hinrichs and S. Ono, 2018. Experimental investigation on the controls of clumped isotopologue and hydrogen isotope ratios in microbial methane. *Geochimica et Cosmochimica Acta* doi:10.1016/j.gca.2018.06.029.
- Gupta, R. S., 2011. Origin of diderm (Gram-negative) bacteria: Antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* 100(2): 171–182. doi:10.1007/s10482-011-9616-8.
- Gyles, C. and P. Boerlin, 2014. Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. *Veterinary pathology* 51(2): 328–40. doi:10.1177/0300985813511131.
- Hafenbradl, D., M. Keller, R. Dirmeier, R. Rachel, P. Roßnagel, S. Burggraf, H. Huber and K. O. Stetter, 1996. *Ferroglobus placidus* gen. nov., sp. nov., a novel hyperthermophilic archaeum that oxidizes Fe²⁺ at neutral pH under anoxic conditions. *Archives of Microbiology* 166(5): 308–314. doi:10.1007/s002030050388.
- Hallam, S. J., N. Putnam, C. M. Preston, J. C. Detter, D. Rokhsar, P. M. Richardson and E. F. DeLong, 2004. Reverse Methanogenesis: Testing the Hypothesis with Environmental Genomics. *Science* 305(5689): 1457–1462. doi:10.1126/science.1100025.
- Hartmann, G. C., E. Santamaria, V. M. Fernández and R. K. Thauer, 1996. Studies on the catalytic mechanism of h₂-forming methylenetetrahydromethanopterin dehydrogenase: para-ortho h₂ conversion rates in h₂ o and d₂ o. *JBIC Journal of Biological Inorganic Chemistry* 1(5): 446–450.
- Hayes, J. M., 2001. Fractionation of Carbon and Hydrogen Isotopes in Biosynthetic Processes. *Reviews in Mineralogy and Geochemistry* 43(1): 225–277. doi:10.2138/gsrmg.43.1.225.
- Heckman, D. S., 2001. Molecular Evidence for the Early Colonization of Land by Fungi and Plants. *Science* 293(5532): 1129–1133. doi:10.1126/science.1061457.
- Hedges, S. B. and S. Kumar, 2003. Genomic clocks and evolutionary timescales. *Trends in Genetics* 19(4): 200–206. doi:10.1016/s0168-9525(03)00053-2.
- Heled, J. and A. J. Drummond, 2015. Calibrated birth-death phylogenetic time-tree priors for Bayesian inference. *Systematic Biology* 64(3): 369–383. doi:10.1093/sysbio/syu089.
- Heuer, V., M. Elvert, S. Tille, M. Krummen, X. P. Mollar, L. R. Hmelo and K.-U. Hinrichs, 2006. Online $\delta^{13}\text{C}$ analysis of volatile fatty acids in sediment/porewater systems by liquid chromatography-isotope ratio mass spectrometry. *Limnology and Oceanography: Methods* 4(10): 346–357. doi:10.4319/lom.2006.4.346.
- Heuer, V. B., J. W. Pohlman, M. E. Torres, M. Elvert and K.-U. Hinrichs, 2009. The stable carbon isotope biogeochemistry of acetate and other dissolved carbon species in deep subseafloor sediments at the northern Cascadia Margin. *Geochimica et Cosmochimica Acta* 73(11): 3323–3336. doi:10.1016/j.gca.2009.03.001.

- Hillis, D. M., T. A. Heath and K. S. John, 2005. Analysis and Visualization of Tree Space. *Syst. Biol* 54(3): 471–482. doi:10.1080/10635150590946961.
- Hinrichs, K. U., 2002. Microbial fixation of methane carbon at 2.7 Ga: Was an anaerobic mechanism possible? *Geochemistry, Geophysics, Geosystems* 3(7). doi:10.1029/2001GC000286.
- Ho, S. Y. W. and S. Duchêne, 2014. Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology* 23(24): 5947–5965. doi:10.1111/mec.12953.
- Holland, G., B. Sherwood Lollar, L. Li, G. Lacrampe-Couloume, G. Slater and C. Ballentine, 2013. Deep fracture fluids isolated in the crust since the precambrian era. *Nature* 497(7449): 357.
- Holler, T., G. Wegener, K. Knittel, A. Boetius, B. Brunner, M. M. M. Kuypers and F. Widdel, 2009. Substantial $^{13}\text{C}/^{12}\text{C}$ and D/H fractionation during anaerobic oxidation of methane by marine consortia enriched in vitro. *Environmental microbiology reports* 1(5): 370–6. doi:10.1111/j.1758-2229.2009.00074.x.
- Holler, T., G. Wegener, H. Niemann, C. Deusner, T. G. Ferdelman, A. Boetius, B. Brunner and F. Widdel, 2011. Carbon and sulfur back flux during anaerobic microbial oxidation of methane and coupled sulfate reduction. *Proceedings of the National Academy of Sciences* 108(52): E1484–E1490.
- Holler, T., G. Wegener, H. Niemann, T. G. Ferdelman, A. Boetius, T. Z. Kristiansen, H. Molina, A. Pandey, J. K. Werner, K. R. Juluri, Y. Xu, D. Glenn, K. Parang and S. H. Snyder, 2012. Correction for Holler et al., Carbon and sulfur back flux during anaerobic microbial oxidation of methane and coupled sulfate reduction. *Proceedings of the National Academy of Sciences* 109(51): 21170–21170. doi:10.1073/pnas.1218683109.
- Holmkvist, L., T. G. Ferdelman and B. B. Jørgensen, 2011. A cryptic sulfur cycle driven by iron in the methane zone of marine sediment (Aarhus Bay, Denmark). *Geochimica et Cosmochimica Acta* 75(12): 3581–3599. doi:10.1016/j.gca.2011.03.033.
- Holtkamp, R., P. Kardol, A. van der Wal, S. C. Dekker, W. H. van der Putten and P. C. de Ruiter, 2008. Soil food web structure during ecosystem development after land abandonment. *Applied Soil Ecology* 39(1): 23–34. doi:10.1016/j.apsoil.2007.11.002.
- Horibe, Y. and H. Craig, 1995. D/h fractionation in the system methane-hydrogen-water. *Geochimica et Cosmochimica Acta* 59(24): 5209–5217.
- Huber, H., M. Thomm, H. König, G. Thies and K. O. Stetter, 1982. *Methanococcus thermolithotrophicus*, a novel thermophilic lithotrophic methanogen. *Archives of Microbiology* 132: 47–50. doi:10.1007/BF00690816.
- Huelsenbeck, J. P., B. Larget, R. E. Miller and F. Ronquist, 2002. Potential Applications and Pitfalls of Bayesian Inference of Phylogeny. *Systematic Biology* 51(5): 673–688. doi:10.1080/10635150290102366.
- Huskey, W. P., 2006. Multiple $\delta^2\text{H}$ isotope probes of hydrogen tunneling. *Hydrogen-Transfer Reactions* pages 1285–1309.

- Husnik, F. and J. P. McCutcheon, 2018. Functional horizontal gene transfer from bacteria to eukaryotes. *Nature Reviews Microbiology* 16(2): 67–79. doi:10.1038/nrmicro.2017.137.
- Hut, G., 1987. Consultants’ group meeting on stable isotope reference samples for geochemical and hydrological investigations. *Report to the Director General* (September 1985): 16–18. doi:18075746.
- Ihrmark, K., N. Asmail, W. Ubhayasekera, P. Melin, J. Stenlid and M. Karlsson, 2010. Comparative molecular evolution of *Trichoderma* chitinases in response to mycoparasitic interactions. *Evolutionary Bioinformatics* 2010(6): 1–26. doi:10.4137/EBO.S4198.
- Ijiri, A., F. Inagaki, Y. Kubo, R. R. Adhikari and S. Hattori, 2018. Deep-biosphere methane production stimulated by geofluids in the Nankai accretionary complex (June): 1–16. doi:10.1126/sciadv.aao4631.
- Inagaki, F., Y. Kubo, M. W. Bowles, V. B. Heuer, A. Ijiri, H. Imachi, M. Ito, M. Kaneko, M. A. Lever, S. Morita, Y. Morono, W. Tanikawa, M. Bihan, S. A. Bowden, M. Elvert, C. Glombitza, D. Gross, G. J. Harrington, T. Hori, K. Li, D. Limmer, M. Murayama, N. Ohkouchi, S. Ono, M. Purkey, Y. Sanada, J. Sauvage, G. Snyder, Y. Takano, E. Tasumi, T. Terada, H. Tomaru, D. T. Wang and Y. Yamada, 2015. Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science* 349(6246): 420–424. doi:10.1126/science.aaa6882.
- Inagaki, F., T. Nunoura, S. Nakagawa, A. Teske, M. Lever, A. Lauer, M. Suzuki, K. Takai, M. Delwiche, F. S. Colwell, K. H. Nealson, K. Horikoshi, S. D’Hondt and B. B. Jorgensen, 2006. Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proceedings of the National Academy of Sciences* 103(8): 2815–2820. doi:10.1073/pnas.0511033103.
- James, T. Y., F. Kauff, C. L. Schoch, P. B. Matheny, V. Hofstetter, C. J. Cox, G. Celio, C. Gueidan, E. Fraker, J. Miadlikowska, H. T. Lumbsch, A. Rauhut, V. Reeb, A. E. Arnold, A. Amtoft, J. E. Stajich, K. Hosaka, G. H. Sung, D. Johnson, B. O’Rourke, M. Crockett, M. Binder, J. M. Curtis, J. C. Slot, Z. Wang, A. W. Wilson, A. Schüßler, J. E. Longcore, K. O’Donnell, S. Mozley-Standridge, D. Porter, P. M. Letcher, M. J. Powell, J. W. Taylor, M. M. White, G. W. Griffith, D. R. Davies, R. A. Humber, J. B. Morton, J. Sugiyama, A. Y. Rossman, J. D. Rogers, D. H. Pfister, D. Hewitt, K. Hansen, S. Hambleton, R. A. Shoemaker, J. Kohlmeyer, B. Volkmann-Kohlmeyer, R. A. Spotts, M. Serdani, P. W. Crous, K. W. Hughes, K. Matsuura, E. Langer, G. Langer, W. A. Untereiner, R. Lücking, B. Büdel, D. M. Geiser, A. Aptroot, P. Diederich, I. Schmitt, M. Schultz, R. Yahr, D. S. Hibbett, F. Lutzoni, D. J. McLaughlin, J. W. Spatafora and R. Vilgalys, 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443(7113): 818–822. doi:10.1038/nature05110.
- Johnson, K. A. and D. E. Johnson, 1995. Methane emissions from cattle. *Journal of animal science* 73(8): 2483–2492.
- Jones, W. J., J. a. Leigh, F. Mayer, C. R. Woese and R. S. Wolfe, 1983. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Archives of Microbiology* 136: 254–261. doi:10.1007/BF00425213.

- Kawagucci, S., M. Kobayashi, S. Hattori, K. Yamada, Y. Ueno, K. Takai and N. Yoshida, 2014. Hydrogen isotope systematics among H₂–H₂O–CH₄ during the growth of the hydrogenotrophic methanogen *Methanothermobacter thermautotrophicus* strain ΔH. *Geochimica et Cosmochimica Acta* 142: 601–614. doi:10.1016/j.gca.2014.07.020.
- Keller, M. D., 1989. Dimethyl sulfide production and marine phytoplankton : the importance of species composition and cell size. *Biological Oceanography* 6(June 2015): 375–382. doi:10.1080/01965581.1988.10749540.
- Kenrick, P., C. H. Wellman, H. Schneider and G. D. Edgecombe, 2012. A timeline for terrestrialization: Consequences for the carbon cycle in the Palaeozoic. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1588): 519–536. doi:10.1098/rstb.2011.0271.
- Kettle, A., M. O. Andreae, D. Amouroux, T. W. Andreae, T. S. Bates, H. Berresheim, H. Bingemer, R. Boniforti, M. a. J. Curran, G. R. DiTullio, G. Helas, G. B. Jones, M. D. Keller, R. P. Kiene, C. Leck, M. Levasseur, G. Malin, M. Maspero, P. Matrai, a. R. McTaggart, N. Mihalopoulos, B. C. Nguyen, A. Novo, J. P. Putaud, S. Rapsomanikis, G. Roberts, G. Schebeske, S. Sharma, R. Simó, R. Staubes, S. Turner and G. Uher, 1999. A global database of sea surface dimethylsulfide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. *Global Biogeochemical Cycles* 13(2): 399. doi:10.1029/1999GB900004.
- Kettles, N. L., S. Kopriva and G. Malin, 2014. Insights into the regulation of DMSP synthesis in the diatom *Thalassiosira pseudonana* through APR activity, proteomics and gene expression analyses on cells acclimating to changes in salinity, light and nitrogen. *PLoS ONE* 9(4). doi:10.1371/journal.pone.0094795.
- Khomyakova, M., Ö. Bükmez, L. K. Thomas, T. J. Erb and I. a. Berg, 2011. A methylaspartate cycle in haloarchaea. *Science* 331(6015): 334–337. doi:10.1126/science.1196544.
- Klapp, S. a., G. Bohrmann, W. F. Kuhs, M. Mangir Murshed, T. Pape, H. Klein, K. S. Techmer, K. U. Heeschen and F. Abegg, 2010. Microstructures of structure I and II gas hydrates from the Gulf of Mexico. *Marine and Petroleum Geology* 27(1): 116–125. doi:10.1016/j.marpetgeo.2009.03.004.
- Klein, A. R., V. M. Fernández and R. K. Thauer, 1995a. H₂–forming n⁵,n¹⁰–methylene-tetrahydromethanopterin dehydrogenase: mechanism of h₂ formation analyzed using hydrogen isotopes. *FEBS letters* 368(2): 203–206.
- Klein, A. R., G. C. Hartmann and R. K. Thauer, 1995b. Hydrogen isotope effects in the reactions catalyzed by h₂–forming n⁵,n¹⁰–methylene-tetrahydromethanopterin dehydrogenase from methanogenic archaea. *The FEBS Journal* 233(1): 372–376.
- Klein, A. R. and R. K. Thauer, 1995. Re–face specificity at c_{14a} of methylene-tetrahydromethanopterin and si–face specificity at c₅ of coenzyme f₄₂₀ for coenzyme f₄₂₀–dependent methylene-tetrahydromethanopterin dehydrogenase from methanogenic archaea. *The FEBS Journal* 227(1–2): 169–174.
- Klinman, J. P. and A. Kohen, 2013. Hydrogen Tunneling Links Protein Dynamics to Enzyme Catalysis. *Annual Review of Biochemistry* 82(1): 471–496. doi:10.1146/annurev-biochem-051710-133623.

- Knoll, A. H., 2014. Paleobiological perspectives on early eukaryotic evolution. *Cold Spring Harbor perspectives in biology* 6(1): 1–14. doi:10.1101/cshperspect.a016121.
- Knoll, A. H., 2017. Biogeochemistry: Food for early animal evolution. *Nature* doi:10.1038/nature23539.
- Knoll, A. H. and M. A. Nowak, 2017. The timetable of evolution. *Science Advances* 3(May): 1–13.
- Kolb, S. and A. Stacheter, 2013. Prerequisites for amplicon pyrosequencing of microbial methanol utilizers in the environment. *Frontiers in Microbiology* 4(SEP): 1–12. doi:10.3389/fmicb.2013.00268.
- Könneke, M., A. E. Bernhard, J. R. De La Torre, C. B. Walker, J. B. Waterbury and D. A. Stahl, 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437(7058): 543–546. doi:10.1038/nature03911.
- Körber, J.-H., H. Sahling, T. Pape, C. dos Santos Ferreira, I. MacDonald and G. Bohrmann, 2014. Natural oil seepage at Kobuleti Ridge, eastern Black Sea. *Marine and Petroleum Geology* 50: 68–82. doi:10.1016/j.marpetgeo.2013.11.007.
- Kröninger, L., J. Gottschling and U. Deppenmeier, 2017. Growth Characteristics of *Methanomassiliicoccus luminyensis* and Expression of Methyltransferase Encoding Genes. *Archaea* 2017. doi:10.1155/2017/2756573.
- Krzycki, J. and W. Kenealy, 1987. Stable carbon isotope fractionation by *Methanosarcina barkeri* during methanogenesis from acetate, methanol, or carbon dioxide-hydrogen. *Applied and ...* 53(10): 2597–2599.
- Kvenvolden, K. A., 1993. Gas hydrates-geological perspective and global change. *Reviews of Geophysics* 31(2): 173–187. doi:10.1029/93RG00268.
- Lacombe-Harvey, M.-È., R. Brzezinski and C. Beaulieu, 2018. Chitinolytic functions in actinobacteria: ecology, enzymes, and evolution. *Applied Microbiology and Biotechnology* doi:10.1007/s00253-018-9149-4.
- Lang, K., J. Schuldes, A. Klingl, A. Poehlein, R. Daniel and A. Brune, 2015. New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of "Candidatus *Methanoplasma termitum*". *Applied and Environmental Microbiology* 81(4): 1338–1352. doi:10.1128/AEM.03389-14.
- Lartillot, N., T. Lepage and S. Blanquart, 2009. PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25(17): 2286–2288. doi:10.1093/bioinformatics/btp368.
- Laso-Pérez, R., G. Wegener, K. Knittel, F. Widdel, K. J. Harding, V. Krukenberg, D. V. Meier, M. Richter, H. E. Tegetmeyer, D. Riedel, H. H. Richnow, L. Adrian, T. Reemtsma, O. J. Lechtenfeld and F. Musat, 2016. Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature* 539(7629): 396–401. doi:10.1038/nature20152.
- Lee, Y. J., I. D. Wagner, M. E. Brice, V. V. Kevbrin, G. L. Mills, C. S. Romanek and J. Wiegel, 2005. *Thermosediminibacter oceani* gen. nov., sp. nov. and *Thermosediminibacter litoriperuensis* sp. nov., new anaerobic thermophilic bacteria isolated from Peru Margin. *Extremophiles* 9(5): 375–383. doi:10.1007/s00792-005-0453-4.

- Lepage, T., D. Bryant, H. Philippe and N. Lartillot, 2007. A General Comparison of Relaxed Molecular Clock Models. *Molecular Biology and Evolution* 24(12): 2669–2680. doi:10.1093/molbev/msm193.
- Lessner, D. J., 2009. Methanogenesis Biochemistry. doi:10.1002/9780470015902.a0000573.pub2.
- Li, C.-Y., T.-D. Wei, S.-H. Zhang, X.-L. Chen, X. Gao, P. Wang, B.-B. Xie, H.-N. Su, Q.-L. Qin, X.-Y. Zhang, J. Yu, H.-H. Zhang, B.-C. Zhou, G.-P. Yang and Y.-Z. Zhang, 2014. Molecular insight into bacterial cleavage of oceanic dimethylsulfoniopropionate into dimethyl sulfide. *Proceedings of the National Academy of Sciences* 111(3): 1026–1031. doi:10.1073/pnas.1312354111.
- Liu, Q. and Y. Liu, 2016. Clumped-isotope signatures at equilibrium of CH₄, NH₃, H₂O, H₂S and SO₂. *Geochimica et Cosmochimica Acta* 175: 252–270. doi:10.1016/j.gca.2015.11.040.
- Lomans, B. P., C. V. D. Drift, a. Pol and H. J. M. O. D. Camp, 2002. Cellular and Molecular Life Sciences Microbial cycling of volatile organic sulfur compounds. *Cellular and molecular life sciences : CMLS* 59: 575–588.
- Lomans, B. P., P. Leijdekkers, J. J. Wesselink, P. Bakkes, A. Pol, C. Van Der Drift and H. J. Op Den Camp, 2001. Obligate Sulfide-Dependent Degradation of Methoxylated Aromatic Compounds and Formation of Methanethiol and Dimethyl Sulfide by a Fresh-water Sediment Isolate, *Parasporobacterium paucivorans* gen. nov., sp. nov. *Applied and Environmental Microbiology* 67(9): 4017–4023. doi:10.1128/AEM.67.9.4017-4023.2001.
- Londry, K. L., K. G. Dawson, H. D. Grover, R. E. Summons and A. S. Bradley, 2008. Stable carbon isotope fractionation between substrates and products of *Methanosarcina barkeri*. *Organic Geochemistry* 39: 608–621. doi:10.1016/j.orggeochem.2008.03.002.
- Lopes, J., L. de Matos, M. Harper, F. Giallongo, J. Oh, D. Gruen, S. Ono, M. Kindermann, S. Duval and A. Hristov, 2016. Effect of 3-nitrooxypropanol on methane and hydrogen emissions, methane isotopic signature, and ruminal fermentation in dairy cows. *Journal of Dairy Science* 99(7): 5335–5344. doi:10.3168/jds.2015-10832.
- Louca, S., P. M. Shih, M. W. Pennell, W. W. Fischer, L. W. Parfrey and M. Doebeli, 2018. Bacterial diversification through geological time. *Nature Ecology & Evolution* page 1. doi:10.1038/s41559-018-0625-0.
- Lozano-Fernandez, J., R. Carton, A. R. Tanner, M. N. Puttick, M. Blaxter, J. Vinther, J. Olesen, G. Giribet, G. D. Edgecombe and D. Pisani, 2016. A molecular palaeobiological exploration of arthropod terrestrialization. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1699): 20150133. doi:10.1098/rstb.2015.0133.
- Lyu, Z., N. Shao, T. Akinyemi and W. B. Whitman, 2018. Methanogenesis. *Current Biology* 28(13): R727–R732. doi:10.1016/j.cub.2018.05.021.
- Ma, Q., S. Wu and Y. Tang, 2008. Formation and abundance of doubly-substituted methane isotopologues (¹³CH₃D) in natural gas systems. *Geochimica et Cosmochimica Acta* 72(22): 5446–5456. doi:10.1016/j.gca.2008.08.014.

- Magallon, S., K. W. Hilu and D. Quandt, 2013. Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *American Journal of Botany* 100(3): 556–573. doi:10.3732/ajb.1200416.
- Magnabosco, C., K. R. Moore, J. M. Wolfe and G. P. Fournier, 2018. Dating phototrophic microbial lineages with reticulate gene histories. *Geobiology* 16(2): 179–189. doi:10.1111/gbi.12273.
- Mamarabadi, M., B. Jensen and M. Lübeck, 2008. Three endochitinase-encoding genes identified in the biocontrol fungus *Clonostachys rosea* are differentially expressed. *Current Genetics* 54(2): 57–70. doi:10.1007/s00294-008-0199-5.
- Marin, J., F. U. Battistuzzi, A. C. Brown and S. B. Hedges, 2017. The Timetree of Prokaryotes: New Insights into Their Evolution and Speciation. *Molecular biology and evolution* 34(2): 437–446. doi:10.1093/molbev/msw245.
- Marlow, J. J., C. T. Skennerton, Z. Li, K. Chourey, R. L. Hettich, C. Pan and V. J. Orphan, 2016. Proteomic Stable Isotope Probing Reveals Biosynthesis Dynamics of Slow Growing Methane Based Microbial Communities. *Frontiers in Microbiology* 7(April): 1–21. doi:10.3389/fmicb.2016.00563.
- McCalley, C. K., B. J. Woodcroft, S. B. Hodgkins, R. a. Wehr, E.-H. Kim, R. Mondav, P. M. Crill, J. P. Chanton, V. I. Rich, G. W. Tyson and S. R. Saleska, 2014. Methane dynamics regulated by microbial community response to permafrost thaw. *Nature* 514(7523): 478–481. doi:10.1038/nature13798.
- McInerney, J. O., J. A. Cotton and D. Pisani, 2008. The prokaryotic tree of life: past, present...and future? *Trends in Ecology and Evolution* 23(5): 276–281. doi:10.1016/j.tree.2008.01.008.
- McNichol, J., S. P. Sylva, F. Thomas, C. D. Taylor, S. M. Sievert and J. S. Seewald, 2016. Assessing Microbial Processes in Deep-Sea Hydrothermal Systems via Incubations at In Situ Temperature and Pressure. *Deep Sea Research Part I: Oceanographic Research Papers* 115: 221–232. doi:10.1016/j.dsr.2016.06.011.
- Mergaert, J., D. Lednická, J. Goris, M. C. Cnockaert, P. De Vos and J. Swings, 2003. Taxonomic study of *Cellvibrio* strains and description of *Cellvibrio ostraviensis* sp. nov., *Cellvibrio fibrivorans* sp. nov. and *Cellvibrio gandavensis* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 53(2): 465–471. doi:10.1099/ij.s.0.02316-0.
- Merzendorfer, H. and L. Zimoch, 2003. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *Journal of Experimental Biology* 206(24): 4393–4412. doi:10.1242/jeb.00709.
- Milkov, A. V., G. E. Claypool, Y.-J. Lee and R. Sassen, 2005. Gas hydrate systems at Hydrate Ridge offshore Oregon inferred from molecular and isotopic properties of hydrate-bound and void gases. *Geochimica et Cosmochimica Acta* 69(4): 1007–1026. doi:10.1016/j.gca.2004.08.021.
- Morrill, P. L., J. G. Kuenen, O. J. Johnson, S. Suzuki, A. Rietze, A. L. Sessions, M. L. Fogel and K. H. Nealson, 2013. Geochemistry and geobiology of a present-day serpentinization site in California: The cedars. *Geochimica et Cosmochimica Acta* 109: 222–240.

- Morris, J. L., M. N. Puttick, J. W. Clark, D. Edwards, P. Kenrick, S. Pressel, C. H. Wellman, Z. Yang, H. Schneider and P. C. J. Donoghue, 2018. The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences* page 201719588. doi:10.1073/pnas.1719588115.
- Munk, A. C., A. Copeland, S. Lucas, A. Lapidus, T. G. Del Rio, K. Barry, J. C. Detter, N. Hammon, S. Israni, S. Pitluck, T. Brettin, D. Bruce, C. Han, R. Tapia, P. Gilna, J. Schmutz, F. Larimer, M. Land, N. C. Kyrpides, K. Mavromatis, P. Richardson, M. Rohde, M. Göker, H.-P. Klenk, Y. Zhang, G. P. Roberts, S. Reslewic and D. C. Schwartz, 2011. Complete genome sequence of *Rhodospirillum rubrum* type strain (S1T). *Standards in Genomic Sciences* 4(3): 293–302. doi:10.4056/sigs.1804360.
- Nascimento, F. F., M. dos Reis and Z. Yang, 2017. A biologist’s guide to Bayesian phylogenetic analysis. *Nature Ecology & Evolution* 1(10): 1446–1454. doi:10.1038/s41559-017-0280-x.
- Nealson, K. H. and P. G. Conrad, 1999. Life: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences* 354(1392): 1923–1939. doi:10.1098/rstb.1999.0532.
- Oakley, T. H., J. M. Wolfe, A. R. Lindgren and A. K. Zaharoff, 2013. Phylotranscriptomics to bring the understudied into the fold: Monophyletic Ostracoda, fossil placement, and pancrustacean phylogeny. *Molecular Biology and Evolution* 30(1): 215–233. doi:10.1093/molbev/mss216.
- Okumura, T., S. Kawagucci, Y. Saito, Y. Matsui, K. Takai and H. Imachi, 2016. Hydrogen and carbon isotope systematics in hydrogenotrophic methanogenesis under H₂-limited and H₂-enriched conditions: implications for the origin of methane and its isotopic diagnosis. *Progress in Earth and Planetary Science* 3(1): 14. doi:10.1186/s40645-016-0088-3.
- Ono, S., D. T. Wang, D. S. Gruen, B. Sherwood Lollar, M. S. Zahniser, B. J. McManus and D. D. Nelson, 2014. Measurement of a Doubly Substituted Methane Isotopologue, ¹³CH₃D, by Tunable Infrared Laser Direct Absorption Spectroscopy. *Analytical Chemistry* 86(13): 6487–6494. doi:10.1021/ac5010579.
- Orcutt, B. N., D. E. Larowe, J. F. Biddle, F. S. Colwell, B. T. Glazer, B. K. Reese, J. B. Kirkpatrick, L. L. Lapham, H. J. Mills, J. B. Sylvan, S. D. Wankel and C. G. Wheat, 2013. Microbial activity in the marine deep biosphere: progress and prospects. *Frontiers in Microbiology* 4(July): 189. doi:10.3389/fmicb.2013.00189.
- Orcutt, B. N., J. B. Sylvan, N. J. Knab and K. J. Edwards, 2011. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiology and molecular biology reviews : MMBR* 75(2): 361–422. doi:10.1128/MMBR.00039-10.
- O’Reilly, J. E., M. dos Reis and P. C. J. Donoghue, 2015. Dating Tips for Divergence-Time Estimation. *Trends in Genetics* 31(11): 637–650. doi:10.1016/j.tig.2015.08.001.
- Panos, A. W. and G. Jones, 2013. Coral animals combat stress with sulphur. *Nature* 502: 634–635. doi:10.1038/nature12698.
- Pape, T., T. Feseker, S. Kasten, D. Fischer and G. Bohrmann, 2011. Distribution and abundance of gas hydrates in near-surface deposits of the Håkon Mosby Mud Volcano,

- SW Barents Sea. *Geochemistry, Geophysics, Geosystems* 12(9): n/a–n/a. doi:10.1029/2011GC003575.
- Pape, T., P. Geprags, S. Hammerschmidt, P. Wintersteller, J. Wei, T. Fleischmann, G. Bohrmann and A. J. Kopf, 2014. Hydrocarbon seepage and its sources at mud volcanoes of the Kumano forearc basin, Nankai Trough subduction zone. *Geochemistry, Geophysics, Geosystems* 18: 1–16. doi:10.1002/2014GC005356. Received.
- Parfrey, L. W., D. J. G. Lahr, A. H. Knoll and L. A. Katz, 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences of the United States of America* 108(33): 13624–9. doi:10.1073/pnas.1110633108.
- Peckmann, J. and V. Thiel, 2004. Carbon cycling at ancient methane-seeps. *Chemical Geology* 205(3–4): 443–467. doi:10.1016/j.chemgeo.2003.12.025.
- Penger, J., R. Conrad and M. Blaser, 2012. Stable carbon isotope fractionation by methylotrophic methanogenic archaea. *Applied and environmental microbiology* 78(21): 7596–602. doi:10.1128/AEM.01773-12.
- Peters, K. E., C. C. Walters and J. M. Moldowan, 2005. *The Biomarker Guide*. Cambridge University Press, Melbourne, 2 edition.
- Peterson, K. J., J. B. Lyons, K. S. Nowak, C. M. Takacs, M. J. Wargo and M. A. McPeck, 2004. Estimating metazoan divergence times with a molecular clock. *Proceedings of the National Academy of Sciences of the United States of America* 101(17): 6536–41. doi:10.1073/pnas.0401670101.
- Philippe, H., Y. Zhou, H. Brinkmann, N. Rodrigue and F. Delsuc, 2005. Heterotachy and long-branch attraction in phylogenetics. *BMC Evolutionary Biology* 5: 1–8. doi:10.1186/1471-2148-5-50.
- Piasecki, A., A. Sessions, M. Lawson, A. A. Ferreira, E. V. S. Neto and J. M. Eiler, 2016. Analysis of the site-specific carbon isotope composition of propane by gas source isotope ratio mass spectrometer. *Geochimica et Cosmochimica Acta* 188: 58–72. doi:10.1016/j.gca.2016.04.048.
- Pine, M. J. and H. Barker, 1956. Studies on the methane fermentation. XII. The pathway of hydrogen in the acetate fermentation. *Journal of bacteriology* 71(6): 644–8.
- Pohlman, J. W., M. Kaneko, V. B. Heuer, R. B. Coffin and M. Whiticar, 2009. Methane sources and production in the northern Cascadia margin gas hydrate system. *Earth and Planetary Science Letters* 287(3–4): 504–512. doi:10.1016/j.epsl.2009.08.037.
- Prieto, M. and M. Wedin, 2013. Dating the Diversification of the Major Lineages of Ascomycota (Fungi). *PLoS ONE* 8(6). doi:10.1371/journal.pone.0065576.
- Quinn, P. K. and T. S. Bates, 2011. The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature* 480(7375): 51–56. doi:10.1038/nature10580.
- Rannala, B. and Z. Yang, 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution* 43(3): 304–311. doi:10.1007/BF02338839.

- Ravenhall, M., N. Škunca, F. Lassalle and C. Dessimoz, 2015. Inferring Horizontal Gene Transfer. *PLoS computational biology* 11(5): e1004095. doi:10.1371/journal.pcbi.1004095.
- Raymann, K., C. Brochier-Armanet and S. Gribaldo, 2015. The two-domain tree of life is linked to a new root for the Archaea. *Proceedings of the National Academy of Sciences of the United States of America* 112(21): 6670–5. doi:10.1073/pnas.1420858112.
- Rees, C., 1973. A steady-state model for sulphur isotope fractionation in bacterial reduction processes. *Geochimica et Cosmochimica Acta* 37(5): 1141–1162. doi:10.1016/0016-7037(73)90052-5.
- Reeves, E. P., J. M. McDermott and J. S. Seewald, 2014. The origin of methanethiol in midocean ridge hydrothermal fluids. *Proceedings of the National Academy of Sciences* page 201400643.
- Reeves, E. P., J. S. Seewald and S. P. Sylva, 2012. Hydrogen isotope exchange between n-alkanes and water under hydrothermal conditions. *Geochimica et Cosmochimica Acta* 77: 582–599. doi:10.1016/j.gca.2011.10.008.
- Reitz, A., T. Pape, M. Haeckel, M. Schmidt, U. Berner, F. Scholz, V. Liebetrau, G. Aloisi, S. M. Weise and K. Wallmann, 2011. Sources of fluids and gases expelled at cold seeps offshore Georgia, eastern Black Sea. *Geochimica et Cosmochimica Acta* 75(11): 3250–3268. doi:10.1016/j.gca.2011.03.018.
- Riedel, M., T. Collett and M. Malone, 2006. Site U1328. In *Proceedings of the Integrated Ocean Drilling Program*, volume 311. doi:10.2204/iodp.proc.311.106.2006.
- Röckmann, T., M. E. Popa, M. C. Krol and M. E. G. Hofmann, 2016. Statistical clumped isotope signatures. *Scientific Reports* 6: 31947. doi:10.1038/srep31947.
- Römer, M., H. Sahling, T. Pape, G. Bohrmann and V. Spieß, 2012. Quantification of gas bubble emissions from submarine hydrocarbon seeps at the Makran continental margin (offshore Pakistan). *Journal of Geophysical Research* 117(C10): C10015. doi:10.1029/2011JC007424.
- Rothman, D. H., G. P. Fournier, K. L. French, E. J. Alm, E. A. Boyle, C. Cao and R. E. Summons, 2014. Methanogenic burst in the end-Permian carbon cycle. *Proceedings of the National Academy of Sciences* 111(15): 5462–5467. doi:10.1073/pnas.1318106111.
- Sahling, H., G. Bohrmann, Y. G. Artemov, A. Bahr, M. Brüning, S. a. Klapp, I. Klaucke, E. Kozlova, A. Nikolovska, T. Pape, A. Reitz and K. Wallmann, 2009. Vodyanitskii mud volcano, Sorokin trough, Black Sea: Geological characterization and quantification of gas bubble streams. *Marine and Petroleum Geology* 26(9): 1799–1811. doi:10.1016/j.marpetgeo.2009.01.010.
- Sahling, H., G. Bohrmann, V. Spiess, J. Bialas, M. Breitzke, M. Ivanov, S. Kasten, S. Krastel and R. Schneider, 2008. Pockmarks in the Northern Congo Fan area, SW Africa: Complex seafloor features shaped by fluid flow. *Marine Geology* 249(3-4): 206–225. doi:10.1016/j.margeo.2007.11.010.
- Sahling, H., M. Römer, T. Pape, B. Bergès, C. dos Santos Fereirra, J. Boelmann, P. Geprägs, M. Tomczyk, N. Nowald, W. Dimmler, L. Schroedter, M. Glockzin and G. Bohrmann,

2014. Gas emissions at the continental margin west of Svalbard: mapping, sampling, and quantification. *Biogeosciences* 11(21): 6029–6046. doi:10.5194/bg-11-6029-2014.
- Saito, M. A., D. M. Sigman and F. M. M. Morel, 2003. The bioinorganic chemistry of the ancient ocean: The co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean-Proterozoic boundary? *Inorganica Chimica Acta* 356: 308–318. doi: 10.1016/S0020-1693(03)00442-0.
- Sanderson, M. J., J. L. Thorne, N. Wikstrom and K. Bremer, 2004. Molecular evidence on plant divergence times. *American Journal of Botany* 91(10): 1656–1665. doi:10.3732/ajb.91.10.1656.
- Santoro, A. E., C. L. Dupont, R. A. Richter, M. T. Craig, P. Carini, M. R. McIlvin, Y. Yang, W. D. Orsi, D. M. Moran and M. A. Saito, 2015. Genomic and proteomic characterization of “*Candidatus Nitrosopelagicus brevis*”: An ammonia-oxidizing archaeon from the open ocean. *Proceedings of the National Academy of Sciences* 112(4): 1173–1178. doi:10.1073/pnas.1416223112.
- Sassen, R., S. Losh, L. Cathles, H. Roberts, J. Whelan, A. Milkov, S. Sweet and D. DeFreitas, 2001. Massive vein-filling gas hydrate: relation to ongoing gas migration from the deep subsurface in the Gulf of Mexico. *Marine and Petroleum Geology* 18(5): 551–560. doi: 10.1016/S0264-8172(01)00014-9.
- Sauer, K. and R. K. Thauer, 1998. Methanol:coenzyme M methyltransferase from *Methanosarcina barkeri*—identification of the active-site histidine in the corrinoid-harboring subunit MtaC by site-directed mutagenesis. *European journal of biochemistry / FEBS* 253(3): 698–705.
- Scheller, S., M. Goenrich, R. Boecher, R. K. Thauer and B. Jaun, 2010. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* 465(7298): 606–8. doi:10.1038/nature09015.
- Scheller, S., M. Goenrich, R. K. Thauer and B. Jaun, 2013. Methyl-Coenzyme M Reductase from Methanogenic Archaea: Isotope Effects on the Formation and Anaerobic Oxidation of Methane. *Journal of the American Chemical Society*.
- Schirrmeister, B. E., J. M. de Vos, A. Antonelli and H. C. Bagheri, 2013. Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. *Proceedings of the National Academy of Sciences* 110(5): 1791–1796. doi:10.1073/pnas.1209927110.
- Schirrmeister, B. E., M. Gugger and P. C. Donoghue, 2015. Corrigendum to Cyanobacteria and the Great Oxidation Event: Evidence from genes and fossils. [Palaeontology. (2015). doi: 10.1111/pala.12178]. *Palaeontology* 58(5): 935–936. doi:10.1111/pala.12193.
- Schleucher, J., C. Griesinger, B. Schwoerer and R. K. Thauer, 1994. H₂-forming n⁵, n¹⁰-methylene-tetrahydromethanopterin dehydrogenase from *Methanobacterium thermoautotrophicum* catalyzes a stereoselective hydride transfer as determined by two-dimensional nmr spectroscopy. *Biochemistry* 33(13): 3986–3993.
- Schoell, M., 1980. The hydrogen and carbon isotopic composition of methane from natural gases of various origins. *Geochimica et Cosmochimica Acta* 44(5): 649–661. doi:10.1016/0016-7037(80)90155-6.

- Schoell, M., 1988. Multiple origins of methane in the earth. *Chemical geology* 71(1-3): 1–10.
- Schwentner, M., D. J. Combosch, J. Pakes Nelson and G. Giribet, 2017. A Phylogenomic Solution to the Origin of Insects by Resolving Crustacean-Hexapod Relationships. *Current Biology* 27(12): 1818–1824.e5. doi:10.1016/j.cub.2017.05.040.
- Schworer, B., V. M. Fernandez, C. Zirngibl and R. K. Thauer, 1993. H₂-forming n₅, n₁₀-methylentetrahydromethanopterin dehydrogenase from methanobacterium thermoautotrophicum. *The FEBS Journal* 212(1): 255–261.
- Sharma, G. and S. Subramanian, 2017. Unravelling the complete genome of Archangium gephyra DSM 2261T and evolutionary insights into myxobacterial chitinases. *Genome Biology and Evolution* 9(5): 1304–1311. doi:10.1093/gbe/evx066.
- Sharpe, S. C., L. Eme, M. W. Brown and A. J. Roger, 2015. Timing the Origins of Multicellular Eukaryotes Through Phylogenomics and Relaxed Molecular Clock Analyses. In I. Ruiz-Trillo and A. M. Nedelcu, editors, *Evolutionary Transitions to Multicellular Life*, pages 3–29. Springer. doi:10.1007/978-94-017-9642-2_1.
- Sherwood Lollar, B., G. Lacrampe-Couloume, K. Voglesonger, T. Onstott, L. Pratt and G. Slater, 2008. Isotopic signatures of CH₄ and higher hydrocarbon gases from Precambrian Shield sites: A model for abiogenic polymerization of hydrocarbons. *Geochimica et Cosmochimica Acta* 72(19): 4778–4795. doi:10.1016/j.gca.2008.07.004.
- Sherwood Lollar, B., T. Onstott, G. Lacrampe-Couloume and C. Ballentine, 2014. The contribution of the precambrian continental lithosphere to global h₂ production. *Nature* 516(7531): 379.
- Sherwood Lollar, B., T. D. Westgate, J. A. Ward, G. F. Slater and G. Lacrampe-Couloume, 2002. Abiogenic formation of alkanes in the Earth's crust as a minor source for global hydrocarbon reservoirs. *Nature* 416(6880): 522–524.
- Shih, P. M., 2015. Photosynthesis and early earth. *Current Biology* 25(19): R855–R859. doi:10.1016/j.cub.2015.04.046.
- Sim, M. S., S. Ono, K. Donovan, S. P. Templer and T. Bosak, 2011. Effect of electron donors on the fractionation of sulfur isotopes by a marine Desulfovibrio sp. *Geochimica et Cosmochimica Acta* 75(15): 4244–4259. doi:10.1016/j.gca.2011.05.021.
- Smrř, J. and V. řatská, 2010. Mycophagous mites and their internal associated bacteria cooperate to digest chitin in soil. *Symbiosis* 52(1): 33–40. doi:10.1007/s13199-010-0099-6.
- Sogin, M. L., M. L. Sogin, H. G. Morrison, H. G. Morrison, J. a. Huber, J. a. Huber, D. Mark Welch, D. Mark Welch, S. M. Huse, S. M. Huse, P. R. Neal, P. R. Neal, J. M. Arrieta, J. M. Arrieta, G. J. Herndl and G. J. Herndl, 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America* 103(32): 12115–20. doi:10.1073/pnas.0605127103.
- Soucy, S. M., J. Huang and J. P. Gogarten, 2015. Horizontal gene transfer: building the web of life. *Nature Reviews Genetics* 16(8): 472–482. doi:10.1038/nrg3962.

- Sousa, D. Z., M. Visser, A. H. Van Gelder, S. Boeren, M. M. Pieterse, M. W. Pinkse, P. D. Verhaert, C. Vogt, S. Franke, S. Kümmel and A. J. Stams, 2018. The deep-subsurface sulfate reducer *Desulfotomaculum kuznetsovii* employs two methanol-degrading pathways. *Nature Communications* 9(1). doi:10.1038/s41467-017-02518-9.
- Spang, A., E. F. Caceres and T. J. Ettema, 2017. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* 357(6351). doi:10.1126/science.aaf3883.
- Spang, A., J. H. Saw, S. L. Jørgensen, K. Zaremba-Niedzwiedzka, J. Martijn, A. E. Lind, R. van Eijk, C. Schleper, L. Guy, T. J. G. Ettema and K. Zaremba, 2015. Europe PMC Funders Group Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521(7551): 173–179. doi:10.1038/nature14447.Complex.
- Srinivasan, R. and H. F. Fisher, 1985. Deuterium isotope effects for the nonenzymic and glutamate dehydrogenase catalyzed reduction of an. alpha.-imino acid by nadh. *Journal of the American Chemical Society* 107(14): 4301–4305.
- Stamatakis, A., 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21): 2688–2690. doi:10.1093/bioinformatics/btl446.
- Stefels, J. and W. H. M. V. Boekel, 1993. L Axenic Cultures of the Marine Phytoplankton. *Marine Ecology Progress Series* 97: 11–18.
- Stefels, J., M. Steinke, S. Turner, G. Malin and S. Belviso, 2007. Environmental constraints on the production and removal of the climatically active gas dimethylsulphide (DMS) and implications for ecosystem modelling. *Phaeocystis, Major Link in the Biogeochemical Cycling of Climate-Relevant Elements* pages 245–275. doi:10.1007/978-1-4020-6214-8_18.
- Stewart, L. C., J.-H. Jung, Y.-T. Kim, S.-W. Kwon, C.-S. Park and J. Holden, 2015. *Methanocaldococcus bathoardescens* sp. nov., a new species of hyperthermophilic methanogen isolated from a volcanically active deep-sea hydrothermal vent. *International Journal of Systematic and Evolutionary Microbiology* pages 0–15. doi:10.1099/ij.s.0.000097.
- Stolper, D., A. Martini, M. Clog, P. Douglas, S. Shusta, D. Valentine, A. Sessions and J. Eiler, 2015. Distinguishing and understanding thermogenic and biogenic sources of methane using multiply substituted isotopologues. *Geochimica et Cosmochimica Acta* 161: 219–247. doi:10.1016/j.gca.2015.04.015.
- Stolper, D., A. Sessions, A. Ferreira, E. Santos Neto, A. Schimmelmann, S. Shusta, D. Valentine and J. Eiler, 2013. Combined ^{13}C –D and D–D clumping in methane: Methods and preliminary results. *Geochimica et Cosmochimica Acta* 126: 169–191. doi:10.1016/j.gca.2013.10.045.
- Stolper, D. A., M. Lawson, C. L. Davis, A. A. Ferreira, E. V. S. Neto, G. S. Ellis, M. D. Lewan, A. M. Martini, Y. Tang, M. Schoell, A. L. Sessions and J. M. Eiler, 2014. Formation temperatures of thermogenic and biogenic methane. *Science* 344(6191): 1500–1503. doi:10.1126/science.1254509.

- Strullu-Derrien, C., A. R. T. Spencer, T. Goral, J. Dee, R. Honegger, P. Kenrick, J. E. Longcore and M. L. Berbee, 2018. New insights into the evolutionary history of Fungi from a 407 Ma Blastocladiomycota fossil showing a complex hyphal thallus. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 373(1739): 20160502. doi:10.1098/rstb.2016.0502.
- Suess, E., M. Torres, G. Bohrmann, R. Collier, J. Greinert, P. Linke, G. Rehder, a. Trehu, K. Wallmann, G. Winckler and E. Zuleger, 1999. Gas hydrate destabilization: enhanced dewatering, benthic material turnover and large methane plumes at the Cascadia convergent margin. *Earth and Planetary Science Letters* 170(1-2): 1–15. doi:10.1016/S0012-821X(99)00092-8.
- Sugimoto, A. and E. Wada, 1993. Carbon isotopic composition of bacterial methane in a soil incubation experiment: Contributions of acetate and CO₂/H₂. *Geochimica et Cosmochimica Acta* 57: 4015–4027.
- Sugimoto, A. and E. Wada, 1995. Hydrogen isotopic composition of bacterial methane: CO₂/H₂ reduction and acetate fermentation 59(7): 1329–1337.
- Sultan, N., G. Bohrmann, L. Ruffine, T. Pape, V. Riboulot, J. Colliat, A. Prunele, B. Denielou, S. Garziglia, T. Himmler, T. Marsset, C. Peters, A. Rabiou and J. Wei, 2014. Pockmark formation and evolution in deep water Nigeria: Rapid hydrate growth versus slow hydrate dissolution. *Journal of Geophysical Research:Solid Earth* 119(December 2011): 2679–2694. doi:10.1029/2010JB007453.Pockmark.
- Summons, R. E., P. D. Franzmann and P. D. Nichols, 1998. Carbon isotopic fractionation associated with methylotrophic methanogenesis. *Organic Geochemistry* 28(7-8): 465–475. doi:10.1016/S0146-6380(98)00011-4.
- Takai, K., K. Nakamura, T. Toki, U. Tsunogai, M. Miyazaki, J. Miyazaki, H. Hirayama, S. Nakagawa, T. Nunoura and K. Horikoshi, 2008. Cell proliferation at 122 c and isotopically heavy ch₄ production by a hyperthermophilic methanogen under high-pressure cultivation. *Proceedings of the National Academy of Sciences* 105(31): 10949–10954.
- Takai, K., K. H. Nealson and K. Horikoshi, 2004. Methanotorris formicicus sp. nov., a novel extremely thermophilic, methane-producing archaeon isolated from a black smoker chimney in the Central Indian Ridge. *International Journal of Systematic and Evolutionary Microbiology* 54(4): 1095–1100. doi:10.1099/ijs.0.02887-0.
- Talamantes, D., N. Biabini, H. Dang, K. Abdoun and R. Berlemont, 2016. Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnology for Biofuels* 9(1): 1–11. doi:10.1186/s13068-016-0538-6.
- Tallant, T. C. and J. a. Krzycki, 1997. Methylthiol:coenzyme M methyltransferase from Methanosarcina barkeri, an enzyme of methanogenesis from dimethylsulfide and methylmercaptopropionate. *Journal of Bacteriology* 179(22): 6902–6911.
- Tallant, T. C., L. Paul and J. A. Krzycki, 2001. The MtsA Subunit of the Methylthiol:Coenzyme M Methyltransferase of Methanosarcina barkeri Catalyses Both Half-reactions of Corrinoid-dependent Dimethylsulfide: Coenzyme M Methyl Transfer. *Journal of Biological Chemistry* 276(6): 4485–4493. doi:10.1074/jbc.M007514200.

- Taylor, J. W. and M. L. Berbee, 2006. Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia* 98(6): 838–849. doi:10.1080/15572536.2006.11832614.
- Thauer, R. K., 1998. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology (Reading, England)* 144: 2377–406. doi:10.1099/00221287-144-9-2377.
- Thauer, R. K., A.-K. Kaster, H. Seedorf, W. Buckel and R. Hedderich, 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature reviews. Microbiology* 6(8): 579–91. doi:10.1038/nrmicro1931.
- Thorne, J. L., H. Kishino and I. S. Painter, 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15(12): 1647–1657. doi:10.1093/oxfordjournals.molbev.a025892.
- Todd, J. D., R. Rogers, Y. G. Li, M. Wexler, P. L. Bond, L. Sun, A. R. J. Curson, G. Malin, M. Steinke and A. W. B. Johnston, 2007. Structural and Regulatory Genes Required to Make the Gas Dimethyl Sulfide in Bacteria. *Science* 315(5812): 666–669. doi:10.1126/science.1135370.
- Tor, J. M. and D. R. Lovley, 2001. Anaerobic degradation of aromatic compounds coupled to Fe(III) reduction by *Ferroglobus placidus*. *Environmental microbiology* 3(4): 281–7.
- Ubhayasekera, W. and M. Karlsson, 2012. Bacterial and fungal chitinase chiJ orthologs evolve under different selective constraints following horizontal gene transfer. *BMC Research Notes* 5. doi:10.1186/1756-0500-5-581.
- Ueno, Y., K. Yamada, N. Yoshida, S. Maruyama and Y. Isozaki, 2006. Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* 440(7083): 516–9. doi:10.1038/nature04584.
- Urey, H. C., 1947. The thermodynamic properties of isotopic substances. *Journal of the Chemical Society (Resumed)* (582): 562. doi:10.1039/jr9470000562.
- Vairavamurthy, A., M. O. Andreae and R. L. Iverson, 1985. Biosynthesis of dimethylsulfide and dimethylpropiothetin by *Hymenomonas carterae* in relation to sulfur source and salinity variations. *Limnology and Oceanography* 30(1): 59–70. doi:10.4319/lo.1985.30.1.0059.
- Valentine, D. L., 2011. Emerging topics in marine methane biogeochemistry. *Annual review of marine science* 3: 147–171.
- Valentine, D. L., A. Chidthaisong, A. Rice, W. S. Reeburgh and S. C. Tyler, 2004. Carbon and hydrogen isotope fractionation by moderately thermophilic methanogens. *Geochimica et Cosmochimica Acta* 68(7): 1571–1590. doi:10.1016/j.gca.2003.10.012.
- Vanwonterghem, I., P. N. Evans, D. H. Parks, P. D. Jensen, B. J. . Woodcroft, P. Hugenholtz and G. W. Tyson, 2016. Methylophilic methanogenesis discovered in the novel archaeal phylum. *Submitted for publication in Nature* doi:10.1038/nmicrobiol.2016.170.
- Varadharajan, C. and H. F. Hemond, 2012. Time-series analysis of high-resolution ebullition fluxes from a stratified, freshwater lake. *Journal of Geophysical Research: Biogeosciences* 117(G2).

- Ver Eecke, H. C., N. H. Akerman, J. A. Huber, D. A. Butterfield and J. F. Holden, 2013. Growth kinetics and energetics of a deep-sea hyperthermophilic methanogen under varying environmental conditions. *Environmental microbiology reports* 5(5): 665–671.
- Vila-Costa, M., J. M. Rinta-Kanto, R. S. Poretsky, S. Sun, R. P. Kiene and M. A. Moran, 2014. Microbial controls on DMSP degradation and DMS formation in the Sargasso Sea. *Biogeochemistry* 120(1-3): 295–305. doi:10.1007/s10533-014-9996-8.
- Vila-Costa, M., R. Simo, H. Harada, J. M. Gasol, D. Slezak and R. P. Kiene, 2006. Dimethylsulfoniopropionate Uptake by Marine Phytoplankton. *Science* 314(5799): 652–654. doi:10.1126/science.1131043.
- Wagner, T., J. Kahnt, U. Ermler and S. Shima, 2016. Didehydroaspartate Modification in Methyl-Coenzyme M Reductase Catalyzing Methane Formation. *Angewandte Chemie - International Edition* 55(36): 10630–10633. doi:10.1002/anie.201603882.
- Waldron, S., J. Lansdown and E. Scott, 1999. The global influence of the hydrogen isotope composition of water on that of bacteriogenic methane from shallow freshwater environments. *Geochimica et Cosmochimica Acta* 63(15): 2237–2245.
- Walker, C. B., J. R. de la Torre, M. G. Klotz, H. Urakawa, N. Pinel, D. J. Arp, C. Brochier-Armanet, P. S. G. Chain, P. P. Chan, A. Gollabgir, J. Hemp, M. Hugler, E. A. Karr, M. Konneke, M. Shin, T. J. Lawton, T. Lowe, W. Martens-Habbena, L. A. Sayavedra-Soto, D. Lang, S. M. Sievert, A. C. Rosenzweig, G. Manning and D. A. Stahl, 2010. Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proceedings of the National Academy of Sciences* 107(19): 8818–8823. doi:10.1073/pnas.0913533107.
- Wang, D. T., D. S. Gruen, B. Sherwood Lollar, K.-u. Hinrichs, L. C. Stewart, J. F. Holden, A. N. Hristov, J. W. Pohlman, P. L. Morrill, M. Könneke, K. B. Delwiche and E. P. Reeves, 2015. Nonequilibrium clumped isotope signals in microbial methane. *Science* 348(6233): 428–431. doi:10.1126/science.aaa4326.
- Wang, D. T., E. P. Reeves, J. M. McDermott, J. S. Seewald and S. Ono, 2018. Clumped isotopologue constraints on the origin of methane at seafloor hot springs. *Geochimica et Cosmochimica Acta* 223: 141–158. doi:10.1016/j.gca.2017.11.030.
- Wang, D. T., P. V. Welander and S. Ono, 2016. Fractionation of the methane isotopologues $^{13}\text{CH}_4$, $^{12}\text{CH}_3\text{D}$, and $^{13}\text{CH}_3\text{D}$ during aerobic oxidation of methane by *Methylococcus capsulatus* (Bath). *Geochimica et Cosmochimica Acta* 192: 186–202. doi:10.1016/j.gca.2016.07.031.
- Wang, F.-P., Y. Zhang, Y. Chen, Y. He, J. Qi, K.-U. Hinrichs, X.-X. Zhang, X. Xiao and N. Boon, 2014. Methanotrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. *The ISME Journal* 8(5): 1069–1078. doi:10.1038/ismej.2013.212.
- Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton, 2009. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9): 1189–1191. doi:10.1093/bioinformatics/btp033.

- Webster, C. R., P. R. Mahaffy, S. K. Atreya, G. J. Flesch, M. a. Mischna, P.-Y. Meslin, K. a. Farley, P. G. Conrad, L. E. Christensen, a. a. Pavlov, J. Martin-Torres, M.-P. Zorzano, T. H. McConnochie, T. Owen, J. L. Eigenbrode, D. P. Glavin, A. Steele, C. a. Malespin, P. D. Archer, B. Sutter, P. Coll, C. Freissinet, C. P. McKay, J. E. Moores, S. P. Schwenzer, J. C. Bridges, R. Navarro-Gonzalez, R. Gellert, M. T. Lemmon and the MSL Science Team, 2014. Mars methane detection and variability at Gale crater. *Science* doi:10.1126/science.1261713.
- Webster, C. R., P. R. Mahaffy, S. K. Atreya, G. J. Flesch, M. A. Mischna, P.-Y. Meslin, K. A. Farley, P. G. Conrad, L. E. Christensen, A. A. Pavlov, J. Martin-Torres, M.-P. Zorzano, T. H. McConnochie, T. Owen, J. L. Eigenbrode, D. P. Glavin, A. Steele, C. A. Malespin, P. D. Archer, B. Sutter, P. Coll, C. Freissinet, C. P. McKay, J. E. Moores, S. P. Schwenzer, J. C. Bridges, R. Navarro-Gonzalez, R. Gellert, M. T. Lemmon and the MSL Science Team, 2015. Mars methane detection and variability at Gale crater. *Science* 347(6220): 415–417. doi:10.1126/science.1261713.
- Wecht, K. J., D. J. Jacob, C. Frankenberg, Z. Jiang and D. R. Blake, 2014. Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data. *Journal of Geophysical Research: Atmospheres* 119(12): 7741–7756. doi:10.1002/2014JD021551.
- Wegener, G., V. Krukenberg, S. E. Ruff, M. Y. Kellermann and K. Knittel, 2016. Metabolic capabilities of microorganisms involved in and associated with the anaerobic oxidation of methane. *Frontiers in Microbiology* 7(FEB): 1–16. doi:10.3389/fmicb.2016.00046.
- Wei, J., T. Pape, N. Sultan, J.-L. Colliat, T. Himmeler, L. Ruffine, A. de Prunel , B. Denielou, S. Garziglia, T. Marsset, C. a. Peters, A. Rabi  and G. Bohrmann, 2015. Gas hydrate distributions in sediments of pockmarks from the Nigerian margin – Results and interpretation from shallow drilling. *Marine and Petroleum Geology* 59: 359–370. doi:10.1016/j.marpetgeo.2014.09.013.
- Welhan, J. and J. Lupton, 1987. Light hydrocarbon gases in guaymas basin hydrothermal fluids: thermogenic versus abiogenic origin. *AAPG Bulletin* 71(2): 215–223.
- Whalen, S., 2005. Biogeochemistry of Methane Exchange between Natural Wetlands and the Atmosphere. *Environmental Engineering Science* 22(1): 73–94. doi:10.1089/ees.2005.22.73.
- Whitehill, A. R., L. M. T. Joelsson, J. A. Schmidt, D. T. Wang, M. S. Johnson and S. Ono, 2017. Clumped isotope effects during OH and Cl oxidation of methane. *Geochimica et Cosmochimica Acta* 196(September): 307–325. doi:10.1016/j.gca.2016.09.012.
- Whiticar, M., 1990. A geochemical perspective of natural gas and atmospheric methane. *Organic Geochemistry* 16: 531–547.
- Whiticar, M. J., 1999. Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chemical Geology* 161(1-3): 291–314. doi:10.1016/S0009-2541(99)00092-3.
- Whiticar, M. J., E. Faber and M. Schoell, 1986. Biogenic methane formation in marine and freshwater environments: Co₂ reduction vs. acetate fermentation— isotope evidence. *Geochimica et Cosmochimica Acta* 50(5): 693–709.

- Wilde, S. A., J. W. Valley, W. H. Peck and C. M. Graham, 2001. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* 409(6817): 175–178. doi:10.1038/35051550.
- Wilson, H. M. and L. I. Anderson, 2004. Morphology and Taxonomy of Paleozoic Millipedes (Diplopoda: Chilognatha: Archipolypoda) From Scotland. *Journal of Paleontology* 78(1): 169–184. doi:10.1666/0022-3360(2004)078<0169:MATOPM>2.0.CO;2.
- Wing, B. A. and I. Halevy, 2014. Intracellular metabolite levels shape sulfur isotope fractionation during microbial sulfate respiration. *Proceedings of the National Academy of Sciences* 111(51): 18116–18125.
- Wolfe, J. M., A. C. Daley, D. A. Legg and G. D. Edgecombe, 2016. Fossil calibrations for the arthropod Tree of Life. *Earth-Science Reviews* 160: 43–110. doi:10.1016/j.earscirev.2016.06.008.
- Wolfe, J. M. and G. P. Fournier, 2018. Horizontal gene transfer constrains the timing of methanogen evolution. *Nature Ecology & Evolution* page 129494. doi:10.1038/s41559-018-0513-7.
- Wongnate, T., D. Sliwa, B. Ginovska, D. Smith, M. W. Wolf, N. Lehnert, S. Raugei and S. W. Ragsdale, 2016. The radical mechanism of biological methane synthesis by methyl-coenzyme m reductase. *Science* 352(6288): 953–958.
- Yanagawa, K., A. Tani, N. Yamamoto, A. Hachikubo, A. Kano, R. Matsumoto and Y. Suzuki, 2016. Biogeochemical Cycle of Methanol in Anoxic Deep-Sea Sediments. *Microbes and environments* 31(2): 190–193. doi:10.1264/jsme2.ME15204.
- Yang, Z. and B. Rannala, 2005. Branch-length prior influences Bayesian posterior probability of phylogeny. *Systematic Biology* 54(3): 455–470. doi:10.1080/10635150590945313.
- Yeung, L. Y., 2016. Combinatorial effects on clumped isotopes and their significance in biogeochemistry. *Geochimica et Cosmochimica Acta* 172: 22–38. doi:10.1016/j.gca.2015.09.020.
- Yoch, D. C., 2002. Dimethylsulfoniopropionate : Its Sources , Role in the Marine Food Web , and Biological Degradation to Dimethylsulfide. *Applied and environmental microbiology* 68(12): 5804–5815. doi:10.1128/AEM.68.12.5804.
- York, D., N. M. Evensen, M. L. Martínez and J. De Basabe Delgado, 2004. Unified equations for the slope, intercept, and standard errors of the best straight line. *American Journal of Physics* 72(3): 367–375.
- Yoshinaga, M. Y., T. Holler, T. Goldhammer, G. Wegener, J. W. Pohlman, B. Brunner, M. M. M. Kuypers, K.-U. Hinrichs and M. Elvert, 2014. Carbon isotope equilibration during sulphate-limited anaerobic oxidation of methane. *Nature Geoscience* 7(3): 190–194. doi:10.1038/ngeo2069.
- Yoshioka, H., S. Sakata and Y. Kamagata, 2008. Hydrogen isotope fractionation by Methanothermobacter thermoautotrophicus in coculture and pure culture conditions. *Geochimica et Cosmochimica Acta* 72(11): 2687–2694. doi:10.1016/j.gca.2008.03.015.

- Yoshioka, Y., T. Tanabe and A. Iguchi, 2017. The presence of genes encoding enzymes that digest carbohydrates in coral genomes and analysis of their activities. *PeerJ* 5: e4087. doi:10.7717/peerj.4087.
- Young, E., I. Kohl, B. Sherwood Lollar, G. Etiope, D. Rumble, S. Li (æĪŒăĴăŒĀ), M. Haghnegahdar, E. Schauble, K. McCain, D. Foustoukos, C. Sutcliffe, O. Warr, C. Ballentine, T. Onstott, H. Hosgormez, A. Neubeck, J. Marques, I. Pérez-Rodríguez, A. Rowe, D. LaRowe, C. Magnabosco, L. Yeung, J. Ash and L. Bryndzia, 2017. The relative abundances of resolved 12CH₂D₂ and 13CH₃D and mechanisms controlling isotopic bond ordering in abiotic and biotic methane gases. *Geochimica et Cosmochimica Acta* 203: 235–264. doi:10.1016/j.gca.2016.12.041.
- Young, E. D., D. Rumble, P. Freedman and M. Mills, 2016. A large-radius high-mass-resolution multiple-collector isotope ratio mass spectrometer for analysis of rare isotopologues of O₂, N₂, CH₄ and other gases. *International Journal of Mass Spectrometry* 401: 1–10. doi:10.1016/j.ijms.2016.01.006.
- Zerkle, A. L., C. H. House and S. L. Brantley, 2005. Biogeochemical Signatures through Time as Inferred from Whole Microbial Genomes. *American Journal of Science* 305: 467–502.
- Zhu, Q., M. Kosoy and K. Dittmar, 2014. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC genomics* 15: 717. doi:10.1186/1471-2164-15-717.
- Zhuang, G. C., F. J. Elling, L. M. Nigro, V. Samarkin, S. B. Joye, A. Teske and K. U. Hinrichs, 2016. Multiple evidence for methylotrophic methanogenesis as the dominant methanogenic pathway in hypersaline sediments from the Orca Basin, Gulf of Mexico. *Geochimica et Cosmochimica Acta* 187: 1–20. doi:10.1016/j.gca.2016.05.005.